University of Alberta

Study of Noncovalent Protein-Carbohydrate Complexes Using Nanoflow Electrospray FT-ICR Mass Spectrometry

by

Weijie Wang

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Chemistry

Edmonton, Alberta

Spring 2005

0-494-08312-3

Library and Archives Canada
Published Heritage Branch
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada
Direction du Patrimoine de l'édition
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN
Our file *Notre reference*
ISBN

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

To realize the full potential of electrospray ionization-mass spectrometry (ES-MS) as a general tool for the determination of protein-ligand binding affinities, a robust experimental protocol is necessary. Using an antibody single-chain fragment (scFv) and its native ligand, αGal[αAbe]αMan, as a model system, the influence of experimental conditions on binding measurements performed with nanoflow electrospray (nanoES) and Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR/MS) were investigated. Mass spectra measured using short spray durations (< 10 min), which minimize pH changes, equimolar analyte concentrations, which minimize the formation of nonspecific complexes, and short accumulation times (< 2 s) in the hexapole of the ion source, which minimize dissociation of the gaseous complexes, accurately reflect the equilibrium distribution of bound and unbound protein in solution. Application of this methodology to the scFv and a series of carbohydrate ligands (αAbe(2-O-CH_3-αMan), βGlc[αAbe]αMan and αGlcαGlcαGal[αAbe]αMan) yields affinities which are in agreement with values obtained by microcalorimetry.

A potential strategy for minimizing nonspecific biological complexes in ES-MS experiments is to selectively dissociate them in the gas phase prior to detection. Blackbody infrared radiative dissociation (BIRD) was performed on the protonated (+10, +11) specific and corresponding nonspecific complexes of scFv with αGal [αAbe]αMan. This study revealed that the nonspecific complex was kinetically more stable at both charge states and, at the +10 charge state, energetically more stable than the specific complex, indicating that selective gas phase dissociation of the nonspecific complex is not feasible.

The mechanism of formation of nonspecific complexes of ubiquitin and carbonic anhydrase with carbohydrates ranging in size from mono- to tetrasaccharides and the nature of their stabilizing interactions were investigated using nanoES-FT-ICR/MS and BIRD. Nonspecific binding was favoured for small (mono- or disaccharide), hydrophilic carbohydrates over larger or hydrophobic carbohydrates, which tend to form gaseous monomer or cluster ions by nanoES; the efficiency was insensitive to the structure of protein and the charge state of the complex. Evidence that both ionic and neutral intermolecular hydrogen bonds stabilize the gaseous complexes was obtained. The use of multivalent metal ions was proposed as a strategy for minimizing the formation of nonspecific complexes of proteins and small molecules.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. John Klassen for his guidance, encouragement and inspiration throughout the course of my graduate study and for teaching me how to write a scientific paper. His dedication and enthusiasm of doing good science will continue to be a great motivation for my career. I would also like to give special thanks to Dr. Elena Kitova for her patience, support and friendship. Without her insightful ideas, my research work could have been much more difficult to be accomplished.

My thanks also go to the members of my research group: Rambod Daneshfar, Igor Sinelnikov, and Jiangxiao Sun. I am grateful to Rambod Daneshfar. He has been always available to help me understand the instrumentation. I thank Igor Sinelnikov for spending the time to help me with numerous computer challenges which I encountered along the way. I also thank Jiangxiao Sun for her participation in some of my research project. The enjoyable and friendly working environment built by group members made my journey over the past five years much more pleasant and at times amusing than would have been the case without such fraternal support and encouragement.

My appreciation is also extended to Professor Paul Kebarle and Dr. Arthur Blades for their very useful suggestions. Many interesting conversations about science and life have broadened my horizons and enriched my experience.

A special thanks goes to Professor David Bundle for generously providing the carbohydrate compounds used in my research work.

Finally, my deepest gratitude goes to my family for their endless unconditional love, moral and financial support from the very beginning.

Table of Contents

Chapter 1 1

Characterization of Noncovalent Biomolecular Complexes by Mass Spectrometry

1.1 Introduction 1

1.2 Mass Spectrometry 5

1.2.1 NanoES-FT-ICR/MS 6

1.2.2 Blackbody Infrared Radiative Dissociation (BIRD) 14

1.3 The Present Work 16

1.4 Literature Cited 21

Chapter 2 28

Influence of Solution and Gas Phase Processes on Protein-Carbohydrate Binding Affinities Determined by Nanoelectrospray Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry

2.1 Introduction 28

2.2 Experimental 33

2.2.1 Protein and Carbohydrate Ligands 33

2.2.2 Mass Spectrometry 33

2.3 Results and Discussion 36

2.3.1 Determining K_{assoc} by MS 36

2.3.2 scFv-αGal[αAbe]αMan Binding Affinity Measurements 38

2.3.3 Influence of Experimental Conditions 45

2.3.3.1 Solution pH 45

2.3.3.2 Analytical Concentration 48

2.3.3.3 In-Source Dissociation 54

2.3.4 Binding Affinities for scFv and Some Related Carbohydrate Ligands 57

2.3.5 Competitive Binding Experiments 57

2.4 Conclusions 61
2.5 Literature Cited 62

Chapter 3 67
Bioactive Recognition Sites May Not Be Energetically Preferred in Protein-Carbohydrate Complexes in the Gas Phase
3.1 Introduction 67
3.2 Experimental 68
3.2.1 Protein and Carbohydrate Ligand 68
3.2.2 Mass Spectrometry 68
3.3 Results and Discussion 70
3.3.1 BIRD Pathways and Kinetics of scFv Nonspecific Complexes 71
3.3.2 Arrhenius Parameters 73
3.3.3 Arrhenius Parameters of CA Nonspecific Complexes 77
3.4 Conclusions 80
3.5 Literature Cited 80

Chapter 4 82
Nonspecific Protein-Carbohydrate Complexes Produced by Nanoelectrospray Ionization. Factors Influencing Their Formation and Stability
4.1 Introduction 82
4.2 Experimental 86
4.2.1 Protein and Carbohydrates 86
4.2.2 Mass Spectrometry 86
4.3 Results and Discussion 89
4.3.1 Formation and Stabilization of Nonspecific Protein-Carbohydrate Complexes 89
4.3.1.1 Influence of Carbohydrate Structure: Size and Hydrophobicity 89

4.3.1.2 Influence of Protein Size and Structure 108
4.3.1.3 Influence of Charge 112
4.3.2 Minimizing Nonspecific Protein-Carbohydrate Complexes 122
4.4 Conclusions 124
4.5 Literature Cited 126

Chapter 5 **131**
Nonspecific Protein-Carbohydrate Complexes Produced by Nanoelectrospray Ionization. Investigating the Nature of the Noncovalent Interactions
5.1 Introduction 131
5.2 Experimental 136
5.2.1 Proteins and Carbohydrates 136
5.2.2 Mass Spectrometry 136
5.3 Results and Discussion 137
5.3.1 Formation of 1:1 and 1:2 Nonspecific Protein-Carbohydrate Complexes 137
5.3.2 Dissociation Pathways 141
5.3.3 Dissociation Kinetics 143
5.3.4 Arrhenius Activation Parameters 156
5.3.4.1 Influence of Carbohydrate Structure 157
5.3.4.2 Influence of Protein Structure 160
5.3.4.3 Influence of Charge State 161
5.4 Conclusions 165
5.5 Literature Cited 166

Chapter 6 **Conclusions and Future Work** **169**
Literature Cited 174

List of Tables:

Table 2.1 Comparison of association constants (K_{assoc}) for scFv and its carbohydrate ligands determined by nanoES-FT-ICR/MS (MS) and isothermal titration calorimetry (ITC). 58

Table 2.2 Association constants (K_{assoc}) for scFv and its carbohydrate ligands determined by nanoES-FT-ICR/MS (MS) and competition experiments performed at two different analyte concentrations (6.1 and 14.8 μM) and by isothermal titration calorimetry (ITC). 61

Table 3.1 Arrhenius parameters determined for the dissociation of gaseous, protonated protein-trisaccharide complexes: $(P + L)^{n+} \rightarrow P^{n+} + L$. 76

Table 5.1 Arrhenius parameters determined for the dissociation of gaseous, protonated and deprotonated protein-trisaccharide complexes: $(P + L)^{n+/z-} \rightarrow P^{n+/z-} + L$, where L = αTal[αAbe]αMan (**3**), αGal[αAbe]αMan (**5**), αGal[αAbe](4-doxyαMan) (**6**), (6-deoxyαGal)[αAbe]αMan (**7**). 157

List of Figures:

Figure 1.1 Schematic representation of ES processes. 7

Figure 1.2 Illustration of the cyclotron motion of a positive ion in the presence of constant magnetic field (*B*). 11

Figure 1.3 Illustration how a mass spectrum is obtained from the ICR motion. 12

Figure 1.4 NanoES-FT-ICR/MS instrument diagram. 14

Figure 2.1 Structures of the oligosaccharide ligands. 34

Figure 2.2 SEM images of two nanoES tips. 35

Figure 2.3 A typical nanoES mass spectrum obtained in positive ion mode for an aqueous solution containing equimolar (6.3 μM) scFv and its native trisaccharide ligand (L) αGal[αAbe]αMan with 1 mM ammonium acetate (pH 7). Insert shows expanded *m/z* region for the +10 charge state; the modified scFv is labeled with an asterisk (*). 39

Figure 2.4 Distribution of R_n obtained from five individual measurements with two different nanoES tips (Tip 1 and Tip 2) for equimolar (6.3 μM) solutions of scFv and αGal[αAbe]αMan: ○, R_{+9}; □, R_{+10}; Δ, R_{+11}; ●, *R* calculated with eq. 2.7. Dashed line indicates the

calculated R value ($\equiv R_{ITC}$) of 0.62. 40

Figure 2.5 Comparison of R_{ave}, measured with five nanoES tips for equimolar (6.3 μM) solutions of scFv and αGal[αAbe]αMan, with the calculated R_{ITC} value. Error bars correspond to values of one standard deviation. Dashed line corresponds to the average R_{ave} value determined from the five sets of measurements (standard deviation is 0.27). 42

Figure 2.6 (a) Plot of R versus spray duration measured for a solution of scFv (14.7 μM) and αGal[αAbe]αMan (14.0 μM). Dashed line indicates the calculated R_{ITC} value of 1.33. (b) Plot of the relative abundance of the $scFv^{n+}$ and $(scFv + \alpha Gal[\alpha Abe]\alpha Man)^{n+}$ ions, at a given charge state, versus spray duration: ●, +9; ■, +10; ○, +11; □, +12. 47

Figure 2.7 Plot of R versus the initial concentration of αGal[αAbe]αMan ($\equiv [L]_o$) measured for a solution containing 18.1 μM scFv: *, MS-derived R_{ave}; ○, calculated R_{ITC}. 49

Figure 2.8 NanoES mass spectra obtained for aqueous solutions containing 18.1 μM scFv and (a) 19.0 μM; (b) 37.9 μM; (c) 50.6 μM αGal[αAbe]αMan (L); (d) equimolar (113 μM) scFv and αGal[αAbe]αMan (L). 50

Figure 2.9 Plot of the dependence of the concentration of αGal[αAbe]αMan at equilibrium (≡ $[L]_{equil}$) on the initial concentration of αGal[αAbe]αMan (≡ $[L]_o$) for cases where $[scFv]_o$ fixed at 18.1 μM, ○ and $[scFv]_o = [L]_o$, ●. 52

Figure 2.10 Plot of $R_{ave}(1+R_{ave})$ versus the initial concentration of αGal[αAbe]αMan (≡ $[L]_o$), measured for equimolar aqueous solutions of scFv and αGal[αAbe]αMan. 53

Figure 2.11 (a) Plot of R versus hexapole accumulation time, obtained from an aqueous solution of scFv (17 μM) and αGal[αAbe]αMan (L) (15 μM). Dashed line corresponds to the calculated R_{ITC} value; (b) Plot of $Ln(I/I_o)$ versus accumulation time, where $I = \sum[I(scFv^{n+})/n]$ and $I_o = \sum[(I(scFv^{n+}) + I(scFv + αGal[αAbe]αMan)^{m+})/n]$ measured with an accumulation time of 1s. 55

Figure 2.12 NanoES mass spectra obtained for equimolar aqueous solutions containing scFv, αGal[αAbe]αMan (L_1), αAbe(2-O-CH_3-αMan) (L_2), and αGlcαGlcαGal[αAbe]αMan (L_3): (a) 14.8 μM; (b) 6.1 μM. 60

Figure 3.1 NanoES mass spectra obtained for buffered aqueous solutions (pH 7.0) containing scFv (19 μM) and its native trisaccharide ligand, αGal[αAbe]αMan (**1**), at (a) 19 μM and (b) 38 μM. 71

Figure 3.2 BIRD mass spectra obtained for the (a) $(scFv + \mathbf{1})^{11+}$, reaction time 1.8 s, reaction temperature149 °C, and (b)

$(scFv + 2(\mathbf{1}))^{11+}$, 1.4 s, 149 °C. 72

Figure 3.3 Plots of the natural logarithm of the normalized abundance (I/I_o) of the gaseous protonated complex $(scFv + \mathbf{1})^{n+}$ ≡ 1:1 (●, +10; ■, +11) and $(scFv + 2(\mathbf{1}))^{n+}$ ≡ 1:2 (○, +10; □, +11) ions versus reaction time measured at 142 °C. The rate constant (k_{obs}) for the loss of **1** from the protonated 1:2 complex was determined from a linear least-square fit. 73

Figure 3.4 Arrhenius plots for the dissociation of the $(scFv + \mathbf{1}_{ns})^{n+}$ (o, +10; □, +11), $(scFv + \mathbf{1}_{sp})^{n+}$ (●, +10; ■, +11) and $(CA + \mathbf{1})^{n+}$ (Δ, +10; ▲, +11) ions. Arrhenius parameters (E_a and A) for each complex were derived from a least-squares linear fit of each Arrhenius plot. 75

Figure 3.5 (a) NanoES mass spectrum obtained for an aqueous solution of CA (8 μM) and **1** (40 μM); (b) BIRD mass spectrum of the $(CA + \mathbf{1})^{10+}$ ion, reaction time 1.5 s, reaction temperature 165 °C. 78

Figure 3.6 Plots of the natural logarithm of the normalized abundance of the protonated complex (a) $(CA + \mathbf{1})^{10+}$ and (b) $(CA + \mathbf{1})^{11+}$ versus reaction time at the temperatures indicated. Kinetic data fit to first-order kinetics. 79

Figure 4.1 Structures of the carbohydrates: D-Gal (**1**), αAbe(2-O-CH_3-αMan) (**2**), αTal[αAbe]αMan (**3**), αAbe(2-O-CH_3-αMan)αGlcβGlc (**4**), D-Glc (**5**) and n-octyl-β-D-glucopyranoside (**6**). 88

Figure 4.2 NanoES mass spectra obtained in positive ion mode for aqueous solutions containing CA (6 μM) and one of carbohydrates (L= **1** - **4**) at (a) 23 μM, (b) 140 μM and (c) 900 μM, with 1 mM ammonium acetate. 90

Figure 4.3 Plot of the fraction of gaseous CA ions engaged in nonspecific binding with L versus [L] measured for solutions containing 6 μM CA and L = **1** (○), **2** (●), **3** (□) **4**, (■). (Fraction of bound protein = 1- ($I_{\Sigma(CA^{n+}/n)}/I_{\Sigma(CA+qL)^{n+}/n}$)). 92

Figure 4.4 Plot of the maximum number of bound carbohydrate ligands (q_{max}) measured from aqueous solutions containing 6 μM protein (CA, □, and Ubq, Δ) and 900 μM L versus the number of saccharide residues: monosaccharide (**1**), disaccharide (**2**), trisaccharide (**3**) and tetrasaccharide (**4**). 93

Figure 4.5 Plot of ratio $q_{ave,exp}/q_{ave,cal}$ versus [L] for solutions of 6 μM CA and L = **1** (○), **2** (●), **3** (□) **4**, (■). $q_{ave,exp}$ corresponds to the average number of bound carbohydrate ligands observed experimentally and $q_{ave,cal}$ is the predicted number based on the concentration ratio [L]/[P]. 95

Figure 4.6 Comparison of experimentally determined distributions (solid bars) of bound carbohydrates in the (CA + qL) ions, formed by nanoES from aqueous solutions containing CA (6 μM) and carbohydrate (L) (a) L = **1** at 23 μM, (b) L =**4** at 23 μM, (c) L = **1** at 140 μM, (d) L = **4** at 140 μM, (e) L = **1** at 900 μM and (f) L = **4** at 900

μM, with 1 mM ammonium acetate, and distributions calculated for a given λ using the Poisson formula, eq. 4.3, (open and shaded bars). 98

Figure 4.7 NanoES mass spectra obtained in positive ion mode for aqueous solutions of CA (6 μM) and 720 μM (a) **5** or (b) **6**, with 1 mM ammonium acetate. 101

Figure 4.8 (a) NanoES mass spectra obtained from the solutions of 6 μM CA and 720 μM **5** or **6**. (b) Corresponding BIRD mass spectra obtained at a cell temperature of 130 °C and a reaction time of 1 s. 102

Figure 4.9 NanoES mass spectrum obtained from an aqueous equimolar (4 μM) solution of **1**, **2**, **3** and **4**, with 1 mM ammonium acetate. 103

Figure 4.10 Formation of nonspecific protein-ligand complexes and ligand clusters as solvent evaporates to dryness during the nanoES process via IEM and CRM. 104

Figure 4.11 NanoES mass spectrum obtained from aqueous solutions containing 6 μM CA and 200 μM of **1** and **4**, with 1 mM ammonium acetate with an accumulation time of (a) 1 s and (b) 2 s in the hexapole. 106

Figure 4.12 (a) NanoES mass spectra obtained from the solutions of 6 μM CA and 300 μM L (**1** - **4**). (b) BIRD mass

spectra obtained at 160 °C and a reaction time of 1 s. 108

Figure 4.13 NanoES mass spectra obtained in positive ion mode for aqueous solutions containing Ubq (6 μM) and one of carbohydrates (L = **1** - **4**) at 900 μM. 110

Figure 4.14 (a) NanoES mass spectra obtained from the solutions of 6 μM CA or Ubq and 100 μM **1**. (b) BIRD mass spectra obtained at 110 °C and a reaction time of 1 s. 112

Figure 4.15 NanoES mass spectra obtained from aqueous solutions containing Ubq (10 μM) and **2** (400 μM) in (a) positive and (b) negative ion modes. 113

Figure 4.16 (a) Isolation mass spectra obtained from the solutions of 6 μM CA and 100 μM **3**; (b) BIRD mass spectra obtained at 167 °C and a reaction time of 2 s. 115

Figure 4.17 (a) Isolation mass spectra obtained from the solutions of 6 μM CA and 180 μM **1**; (b) BIRD mass spectra obtained at 150 °C and a reaction time of 1 s. 116

Figure 4.18 (a) Isolation mass spectra obtained from the solutions of 6 μM Ubq and 80 μM **3**; (b) BIRD mass spectra obtained at 160 °C and a reaction time of 1.5 s. 117

Figure 4.19 NanoES mass spectra obtained from the solutions of 10 μM Ubq and 100 μM **1** in negative ion mode (a) and in positive

ion mode (c) and the corresponding BIRD mass spectra obtained at 91 °C and a reaction time of 2 s (b and d). 118

Figure 4.20 NanoES mass spectra obtained from the solutions of 10 μM Ubq and 170 μM **2** in negative ion mode (a) and in positive ion mode (c) and the corresponding BIRD mass spectra obtained at 104 °C and a reaction time of 2 s (b and d). 119

Figure 4.21 NanoES mass spectra obtained from the solutions of 10 μM Ubq and 100 μM **3** in negative ion mode (a) and in positive ion mode (c) and the corresponding BIRD mass spectra obtained at 159 °C and a reaction time of 3 s (b and d). 120

Figure 4.22 NanoES mass spectra obtained from the solutions of 10 μM Ubq and 100 μM **4** in negative ion mode (a) and in positive ion mode (c) and the corresponding BIRD mass spectra obtained at 185 °C and a reaction time of 3 s (b and d). 121

Figure 4.23 NanoES mass spectra obtained from the solutions of 6 μM CA, 500 μM **2** and 1 mM $CH_3COO\ NH_4$ in the absence (a) and presence of 40 μM salt: NaCl (b), $CaCl_2$ (c) or $MgCl_2$ (d). 123

Figure 5.1 Structures of the carbohydrates: D-Gal (**1**), αAbe(2-O-CH_3-αMan) (**2**), αTal[αAbe]αMan (**3**), αAbe(2-O-CH_3-αMan)αGlcβGlc (**4**), αGal[αAbe]αMan (**5**), αGal[αAbe](4-deoxyαMan) (**6**), (6-deoxyαGal)[αAbe]αMan (**7**) 138

Figure 5.2 NanoES mass spectra obtained in the positive ion

mode from solutions of protein and **3**, (a) 10 μM CA and 47 μM **3**; (b) 10 μM CA and 71 μM **3**; (c) 10 μM CA and 47 μM **3** with 0.1 mM imidazole; (d) 10 μM Ubq and 45 μM **3**; (f) 10 μM BPTI and 63 μM **3**; (e) nanoES mass spectrum obtained in the negative ion mode from a solution of 10 μM Ubq and 41 μM **3**. 139

Figure 5.3 BIRD mass spectra obtained for: (a) $(CA + \mathbf{1})^{10+}$, at a temperature of 100 °C and a reaction time 1.5s; (b) $(CA + \mathbf{2})^{10+}$, 128 °C and 1.5s; (c) $(CA + \mathbf{3})^{10+}$, 171 °C and 1.5s; (d) $(CA + \mathbf{4})^{10+}$, 187 °C and 6s; (e) $(CA + (2)\mathbf{3})^{10+}$, 171 °C and 1.5s; (f) $(Ubq + \mathbf{3})^{5-}$, 163 °C and 1.5s. 142

Figure 5.4 Plots of the natural logarithm of the normalized abundance of the complex ion, $(CA + L)^{10+}$, where L = **1** – **4**, versus reaction time at the temperatures indicated: (a) ○, L = **3**; ■, L = **4**; (kinetic data fit to first order kinetics) (b) ○, L = **1**; ■, L = **2**. 144

Figure 5.5 Plots of the natural logarithm of the normalized abundance of the complex ion $(Ubq + L)^{5+}$, where L = **1** – **4**, versus reaction time at the temperatures indicated: (a) ○, L = **3**; ■, L = **4**; (kinetic data fit to first order kinetics) (b) ○, L = **1**; ■, L = **2**. 145

Figure 5.6 Plots of the natural logarithm of the normalized abundance of the complex ion: (a) $(CA + 2(\mathbf{3}))^{10+}$, (b) $(CA + 2(\mathbf{3}))^{11+}$ and (c) $(Ubq + 2(\mathbf{3}))^{5+}$, at the temperatures indicated. Kinetic data fit to first order kinetics) 150

Figure 5.7 The ratio of the rate constant of the dissociation of 1:2 complex ion (k_{obs}) and the rate constant for the dissociation of corresponding 1:1 complex ion ($k_{1:1}$) at a given temperature (■, $(CA + (2)\mathbf{3})^{10+}$; ●, $(CA + (2)\mathbf{3})^{11+}$ and ▲, $(Ubq + (2)\mathbf{3})^{5+}$). 152

Figure 5.8 Comparisons of the experimental breakdown curves derived from the kinetic data for the complexes: (a) $(CA + 2(\mathbf{3}))^{10+}$ at 139 °C; (b) $(CA + 2(\mathbf{3}))^{11+}$ at 150 °C; (c) $(Ubq + 2(\mathbf{3}))^{5+}$ at 119 °C, and the corresponding theoretical breakdown curves calculated using the kinetic *model I* ($k_1 = k_1' = k_{1:1}$ and $k_2 = k_2' = k_{obs} - k_{1:1}$ (—)); kinetic *model II* ($k_1 = k_2 = 1/2k_{obs}$ and $k_1' = k_2' = k_{1:1}$ (·····));kinetic *model III* ($k_1 = k_2 = k_1' = k_2' = 1/2k_{obs}$ (— ·— ·)). 154

Figure 5.9 Arrhenius plots for the loss of L from the protonated complex ions, $(CA + L)^{n+}$ where n is (a) 10 and (b) 11, (○: L = **5**; ▲: L = **3**; ●: L = **6**; □: L = **7**). 159

Figure 5.10 Arrhenius plots for the loss of **3** from the protonated complex ions, $(CA + \mathbf{3})^{n+}$: □ (n = 8), ■ (9), ○ (10), ● (11), Δ (12). 162

Figure 5.11 Arrhenius plots for the loss of L from the protonated and deprotonated complex ions (○, $(Ubq + \mathbf{5})^{5+}$; ▲, $(Ubq + \mathbf{3})^{5+}$;■, $(Ubq + \mathbf{3})^{5-}$ and □, $(BPTI + \mathbf{5})^{5+}$). 164

List of Schemes:

Scheme 5.1 Dissociation pathways of protein-ligand complex ions, $(P + L)^{n+}$, which consists of multiple conformers with distinct dissociation rate constants. 147

Scheme 5.2 Dissociation pathways of protein-ligand complex, $(P + L)^{n+}$ which undergoes conformational changes over the course of the reaction. 147

Scheme 5.3 Sequential dissociation pathways of a 1:2 protein-carbohydrate complex $(P + L_1 + L_2)^{n+}$. L_1 and L_2 correspond to the same carbohydrate bound at two different sites. 149

List of Abbreviations

BIRD	Blackbody infrared radiative dissociation
BPTI	Bovine pancreatic trypsin inhibitor
CA	Bovine carbonic anhydrase II
CID	Collision-induced dissociation
Con A	Apo-concanavalin
CRM	Charge residue model
EDTA	Ethylenediaminetetraacetic acid
ES	Electrospray
FAC/MS	Frontal affinity chromatography–mass spectrometry
FT-ICR/MS	Fourier transform ion cyclotron resonance mass spectrometry
IEM	Ion evaporation model
ITC	Isothermal titration calorimetry
MALDI	Matrix-Assisted Laser Desorption Ionization
MS	Mass spectrometry
m/z	Mass-to-charge ratio
nanoES	Nanoflow electrospray
NMR	Nuclear magnetic resonance
Pt	Platinum
REX	Rapid-exchange limit
rf	Radio frequency
scFv	Single-chain variable domain fragment
SEM	Scanning electron microscopy

S/N	Signal-to-noise ratio
SPR	Surface plasma resonance
TOF	Time of flight
Ubq	Ubiquitin

Chapter 1

Characterization of Noncovalent Biomolecular Complexes by Mass Spectrometry

1.1 Introduction

Noncovalent biomolecular complexes between proteins, small molecules, carbohydrates, DNA, RNA, and metal ions play a central role in many important biological processes, such as gene transcription, cell signaling, immune response and ion transport [1]. How the components of these complexes are bound influences the functioning of living cells and is directly related to health and disease. Since biochemical function is typically mediated by the inter- and intramolecular interactions present in biomolecules, understanding the links between structure and function of biological complexes provides an approach to understanding the origin of disease and the effects of therapy at the molecular and cellular level.

To achieve this understanding, a multi-disciplinary approach that combines molecular biology, chemistry, biophysical chemistry and physics is required due to the structural complexity of biomolecules. At present, a variety of different methodologies and techniques utilizing optical spectroscopy (e.g. UV absorption, circular dichroism, vibrational spectroscopy and fluorescence), nuclear magnetic resonance (NMR), light scattering, differential scanning calorimetry, isothermal titration calorimetry (ITC), surface plasma resonance (SPR) and X-ray crystallography are widely used to characterize noncovalent complexes, including identification of the components of a complex, mapping the interaction interface, determination of the secondary, tertiary or

quaternary structure of a complex, and probing the mechanisms governing the molecular recognition [2].

Mass spectrometry (MS) has been widely used in biomedical research as it offers advantages in sensitivity, speed, specificity and accuracy of mass determination [3]. The advent of soft ionization methods, in particular, electrospray (ES) [4], and nanoflow electrospray (nanoES) [5], have led to a remarkable revolution in the application of MS in studying the complexation processes of macromolecules in solution. Beyond their ability to transfer non-volatile macromolecules from aqueous solutions to the gas phase in an ionized form, ES and nanoES are able to preserve intact noncovalent complexes in the gas phase. Thus, it becomes possible to investigate noncovalent biomolecular complexes by using ES-MS. The first observations of the FKBP-FK506 receptor-ligand complex [6] and the noncovalent binding of heme to myoglobin [7] obtained by electrospraying aqueous solutions at near neutral (or physiological) pH demonstrated the feasibility of using ES-MS for studying complexes that associate through specific noncovalent interactions in solution. Since these initial reports, a variety of noncovalent complexes have been studied, including enzyme-substrate, protein-cofactor, antibody-antigen, protein-DNA complexes, protein assemblies, oligonucleotide duplexes and DNA-small molecules [8-21]. A direct mass measurement of noncovalent complexes larger than 2 MDa has been reported by Robinson and coworkers recently [22]. A detailed overview of the observation of noncovalent biomolecular complexes by ES-MS has appeared in several review articles [23].

Besides the detection of these complexes, the binding stoichiometry of the complexes can be easily deduced from the molecular weight measurement [24]. In

addition, the relative [25] and absolute binding affinities [26] of the complexes can be evaluated from relative ion abundances observed in the mass spectra. In many cases, the binding information obtained from these gas phase-based measurements is consistent with the expected solution binding characteristics. Recently, using a variable-temperature nanoES-MS, the evaluation of thermal stability of protein assemblies in solution and the determination of the association enthalpies and entropies for protein-ligand complexes have been demonstrated [27]. Taking advantage of its ability to perform simultaneous measurements on mixtures of analytes, ES-MS is an extremely useful tool for screening possible drug targets from combinatorial libraries in a rapid, accurate, and sensitive manner [28a, b]. As reported by Wigger and coworkers [28b], ES combined with Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR/MS) and IR multiphoton dissociation has been applied to detect intact protein-ligand complexes from a solution containing a protein and a combinatorial library of 324 peptides. Following detection of the intact noncovalent complexes, selective dissociation via gas phase methods was used to identify the ligands bound to the protein.

Used in conjunction with gas phase dissociation techniques, ES-MS is also being employed for the determination of the composition of complexes [29], the investigation of intrinsic noncovalent interactions [30] and the quaternary structure of biological complexes [31]. Evidence from gas-phase studies of biomolecular complexes suggests that to some extent, the ES-MS observations of gaseous noncovalent complexes reflect the nature of the interactions found in solution [32]. However, in some cases, the stabilities of complexes obtained from gas phase measurements do not correlate well with those obtained in solution [33]. The nature of the noncovalent interactions may have an

effect on the success of the mass spectrometry experiment. During the transfer from solution to the solvent-free gas phase, electrostatic interactions and hydrogen bonds may be strengthened. Some examples demonstrate that the contributions from electrostatic interactions are more significant than hydrophobically driven interactions for the preservation of gas phase noncovalent complexes [33a-c]. Therefore, caution is required in the use of gas phase results from ES-MS experiments to predict the binding strength of noncovalent interactions formed in the solution phase. If conditions can be found where the gaseous complex ions retain the specific interactions that allow the partners to recognize each other in solution, mass spectrometry may find widespread application to identify the binding site and measure the binding strength of specific noncovalent interaction. Mapping the intermolecular interactions present in gaseous protein-ligand complexes can be realized by applying thermal dissociation methods combined with a functional group replacement strategy. For example, with proteins, amino acid residues that are involved in solution binding can be altered by site-directed mutagenesis to change the solution binding. A comparison of the relative binding energy of the various complexes, measured in the gas phase, may determine if the solution interaction is preserved in the gas phase. Using this approach, the first evidence that the specific interaction is preserved in the gaseous protein complex was demonstrated by Klassen and coworkers [30f]. In summary, ES-MS has emerged as an invaluable tool to investigate the biochemical properties of noncovalent complexes, although the methodology needs to be further developed.

In the following sections, fundamental aspects of the MS techniques used in my research, nanoES-FT-ICR/MS and the time-resolved thermal dissociation method,

blackbody infrared radiative dissociation (BIRD), will be briefly described. Also a brief introduction of my research work will be presented.

1.2 Mass Spectrometry

Mass spectrometric analyses involve the formation of gas phase ions by ionization methods and detection of the mass-to-charge ratio (*m/z*) of these ions by mass analyzers. The choice of ionization method depends on the nature of the sample and the type of information required from the analysis. In general, ionization methods can be classified into six categories. (1) Low pressure gas-phase ionization technique, including electron impact ionization and chemical ionization, used for small volatile compounds. (2) Field desorption and ionization techniques, which tend to produce little or no fragmentation for non-volatile molecules with moderate molecular weight. (3) Particle bombardment techniques, including fast atom bombardment and secondary ion mass spectrometry, used for non-volatile analytes soluble in a liquid matrix. (4) Matrix-assisted laser desorption/ionization (MALDI), used for the ionization of non-volatile macromolecules. (5) Atmospheric pressure ionization including ES, nanoES and atmospheric pressure chemical ionization. ES and MALDI, which are often referred to as "soft ionization" techniques, usually generate intact macromolecular ions and are, therefore, well suited for macromolecular mass determination. In MALDI, analytes are co-crystalized with a chemical matrix under vacuum. Upon exposure to pulsed laser radiation, the matrix sublimates and carries the analyte molecules into the gas phase. In ES, analytes are sprayed directly from solution by a strong electric field gradient at atmospheric pressure. ES is more frequently used for mass spectrometric investigation of noncovalent

complexes than MALDI, due to its ability to maintain the complexes in physiological conditions prior to ionization.

Mass analyzers include magnetic sector, quadrupole, ion trap, time of flight (TOF), and FT-ICR. With high mass resolution and wide mass detection range, TOF and FT-ICR are able to determine the stoichiometry of the binding partners, even for small molecules (e.g., inhibitors and drug molecules) bound to larger molecular mass targets. MALDI typically produces ions in pulsed intervals under vacuum, making it most suitable for coupling with ion trap, TOF and FT-ICR. ES produces ions continuously, making it amenable to coupling with scanning mass analyzers such as quadrupole and magnetic sector mass analyzer. The versatility of MS techniques is continuously enhanced through the combination of multiple mass analyzers.

1.2.1 NanoES-FT-ICR/MS

The mechanism of nanoES, a low solution flow version of conventional ES, invented by Matthias Mann in 1994 [5], is similar to that of ES. As described by Kebarle and coworkers [34], ES processes involves the production of charged droplets from electrolyte dissolved in a solvent; shrinkage of the droplets by solvent evaporation accompanied with droplet fissions, formation of very small, highly charged droplets from which gaseous ions are produced.

Shown in Figure 1.1 [34] is a diagram describing the ES processes. The high positive voltage applied to the capillary induces charge separation of electrolytes in solution, positive charges drift towards the liquid surface leading to the formation of a liquid cone referred to as a Taylor cone. At a sufficiently high electric field, the liquid

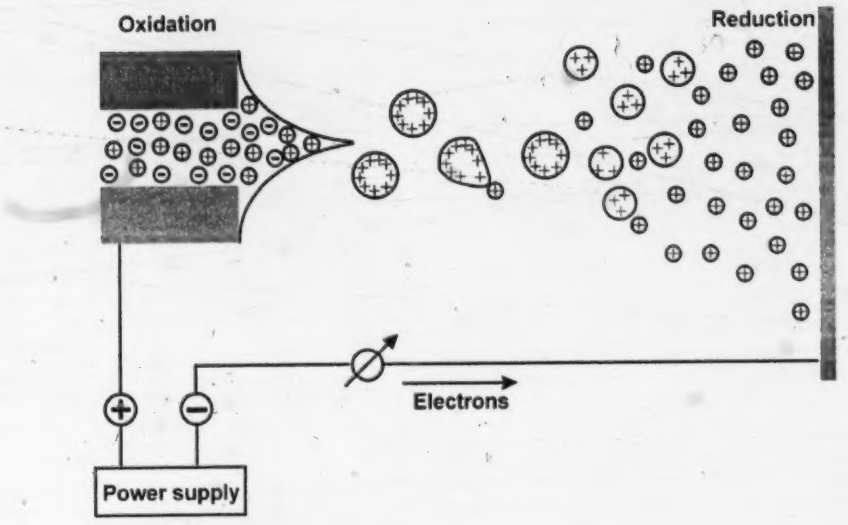


Figure 1.1 Schematic representation of ES processes.

cone becomes unstable and emits a thin liquid filament which subsequently breaks up into small positively charged droplets. With solvent evaporation, these droplets start to shrink into the smaller droplets. As the charge density on the droplet surface increases to the Rayleigh limit (the point at which the Coulombic repulsion of the surface charges is equal to the surface tension of the droplets), these droplets undergo Rayleigh fission, eventually forming small highly charged offspring droplets. Production of gas phase ions from small droplets follows one of two mechanisms: the ion evaporation model and the charge residue model. The ion evaporation model, proposed by Iribarne and Thomson [35] assumes ion emission directly from very small and highly charged droplets. This model accounts for the production of gas phase ions of small molecules. In the charge residue model, proposed by Dole [36], the droplets undergo many fissions, finally producing droplets containing a few analyte molecules which, by further solvent

evaporation, results in gaseous ions. This model is believed to be most significant for macromolecules.

According to the ES processes described above, the question of whether the solution binding is preserved following ES conversion to the gas phase is raised. To better understand this, several issues should be considered. First, during the ion separation in solution, electrochemical reactions (i.e. solvent oxidation. emitter electrode corrosion and analyte oxidation) occur at the electrode in the ES tip to supply excess charges required to maintain the productions of charged droplets. As a result, the composition (pH, metal ion concentration and analyte charge state) of the initial solution may be altered during the course of sampling. It has been reported by Van Berkel and coworkers [37] that the degree of alteration of solution composition by electrochemical reactions depends on the experimental conditions (solvent, ES currents, flow rates, electrode materials, *etc.*) and analyte properties (pKa, metal-binding affinity, equilibrium redox potential, *etc.*). Since most biocomplexes are sensitive to the solution pH, one should be aware of the influence of electrochemical reactions occurring during the ES processes. The second factor to be considered is whether the distribution of species at equilibrium in solution could be perturbed in the ES droplets due to the increasing ionic strength and analyte concentration during the solvent evaporation. A comparison of the residence time of droplets in the ES source and the lifetime of the complexes, determined by the kinetics of association/dissociation reactions, provides useful information about the possibility of an equilibrium shift during the ES process. However, due to the complexity of the ES processes, estimates of droplet lifetimes are uncertain. Factors such as the size of droplets, fission events, solvent evaporation rates and temperatures of the

droplets will influence the lifetime of droplets. In spite of these uncertainties, an insightful study on the complexation of strontium with EDTA by Wang *et al.* [38] has shown that a smaller than expected shift in the equilibrium distribution, based on the kinetics of reaction and the estimated time scale of the ES process. This study suggests that kinetically labile species could be monitored by ES. The third consideration is related to the issue of ionization efficiency. The binding affinity of complexes is evaluated from the relative abundances of ions corresponding to solution species observed in the mass spectra. Any difference in the ionization efficiency would lead to incorrect relative ion abundances in the mass spectra. The ionization efficiency is dependent on the chemical properties of analytes, solvents, charge agents, and droplet size. To eliminate the influence of ionization efficiency on the binding constant measurements, ions with similar ionization efficiencies are chosen to derive the binding constant. If this is not the case, the equilibrium concentration of each component in solution can be derived from the intensity of the corresponding ion in the mass spectra, based on a calibration plot of ion abundance as a function of solution concentration. This approach relies on the assumption that the ES response to a particular compound is dependent only on its concentration in solution and is not affected significantly by the presence of other species. This approach has been successfully applied to the determination of binding constants of vancomycin and ristocetin complexes with peptides [39]. The last consideration is the formation of complexes from random aggregation during the ES process. These artificial complexes, which are commonly referred to as "nonspecific complexes", may be sufficiently long-lived in the gas phase to be detected. Their presence in the mass spectra may obscure the binding stoichiometry of specific

complexes in solution. All these issues represent challenges in developing an ES-MS-based methodology for the characterization of solution phase noncovalent complexes.

Instead of conventional ES, nanoES is used in the present work. It operates at lower solution flow rates, typically 10-50 nL/min and emits correspondingly smaller droplets than conventional ES. The low flow rate of nanoES was realized experimentally with nanoES tips pulled from glass capillaries to a fine tip with a diameter in the order of μm. In our laboratory, aluminosilicate glass capillaries, pulled with a laser micropipette puller to a very fine tip with an external diameter around 2-7 μm were used. The features of nanoES make it more convenient than conventional ES for investigating noncovalent complexes. Only picomoles or less of analyte per analysis are needed in using nanoES, a very important feature in the analysis of limited amounts of biosamples. In addition to being more sensitive than ES, nanoES readily allows for the transfer of noncovalent complexes from buffered aqueous solutions to the gas phase and, therefore, can be directly performed on the complex solution remained under near physiological conditions. The short lifetimes of the nanoES droplets, which are estimated to be tens of microseconds, are likely advantageous for preserving the original solution composition throughout the formation of gaseous ions. Furthermore, nanoES can minimize the nonspecific aggregation from ES processes due to the presence of only a few analyte molecules for each droplet.

After ions are generated by nanoES, FT-ICR/MS was used for the mass detection. Unlike other mass analyzers, FT-ICR/MS is a technique that converts *m/z* to an experimentally measurable ion cyclotron orbital frequency [40]. Shown in Figure 1.2 is

the cyclotron motion of an ion with mass, m, and charge, z, in a uniform magnetic field, B. The ion moves in a circular orbit perpendicular to the magnetic field direction.

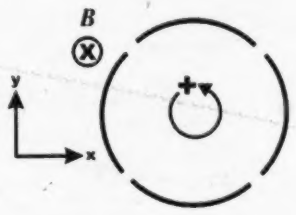


Figure 1.2 Illustration of the cyclotron motion of a positive ion in the presence of constant magnetic field (B).

The cyclotron frequency, ω_c is expressed in eq. 1.1, in which B is in tesla, m is in kilogram, and z is in multiples of the elementary charge, e.

$$\omega_c = \frac{zeB}{m} \tag{1.1}$$

The cyclotron frequency of an ion is inversely proportional to its m/z and directly proportional to the strength of the applied magnetic field. Ions with lower m/z have higher cyclotron frequencies. A notable feature of eq. 1.1 is that all ions of a given m/z rotate at the same frequency, independent of velocity. This property makes ICR especially amenable to MS, because ion frequency is insensitive to kinetic energy, so that focusing of translational energy is not essential for the precise determination of m/z.

Although ions in a static magnetic field move in cyclotron orbits, they will not generate any signal if placed between a pair of detection electrodes. Figure 1.3 demonstrates how a mass spectrum is generated from the ion motion. In order to produce

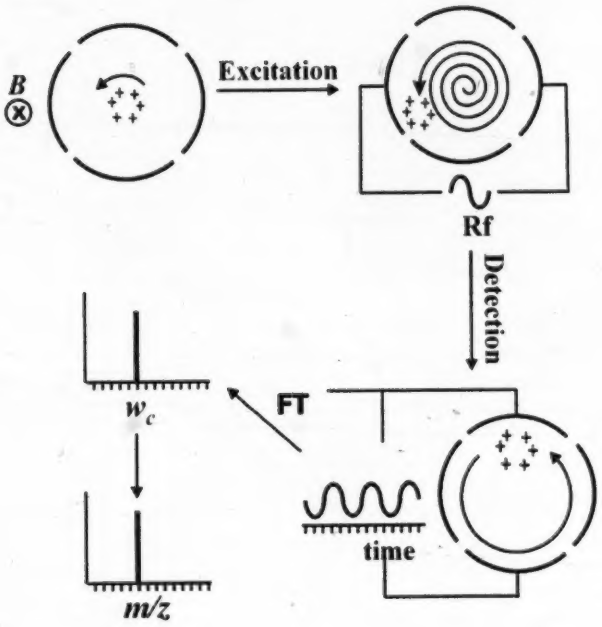


Figure 1.3 Illustration how a mass spectrum is obtained from the ICR motion.

a signal, a packet of ions of a given *m/z* needs to be excited by applying an oscillating electrical field such as provided by an ac signal generator. If the frequency of the applied field is the same as the ω_c of the ions, the ions absorb energy thus increasing their orbital radius but keeping a constant cyclotron frequency. This phenomenon provides the basis for ion cyclotron resonance mass spectrometry because ions having a different cyclotron frequency are not accelerated. To detect ions, a differential image current between two opposite detection plates induced by a coherently orbiting ion packet induces is required. How is the image current obtained? As the ion(s) in a circular orbit approach the top plate, electrons are attracted to this plate from ground. Then as the ion(s) circulate towards the bottom plate, the electrons travel back down to the bottom plate. This motion

of electrons moving back and forth between the two plates produces a detectable current. The amplitude of this current is proportional to the number of ions in the analyzer cell while its frequency is the same as the cyclotron frequency of the ions. A small ac voltage is generated across a resistor and this signal is amplified and detected. The ions are therefore detected without ever colliding with the electrodes. This non-destructive detection scheme is unique to FT-ICR and allows for improved sensitivity and versatility compared to more traditional approaches that utilize destructive detection methods. The detected image current is transformed from the time domain into a frequency domain signal by a Fourier transform. Because the cyclotron frequency is related to *m/z*, a mass spectrum can be obtained. Cyclotron frequencies can be measured with very high precision, leading to high accuracy mass measurements and ultra-high resolving power.

Shown in Figure 1.4 is a simplified diagram of the nanoES-FT-ICR/MS used in the present work. A buffered aqueous solution containing analyte molecules is loaded into a nanoES tip by syringe and the solution is sprayed at atmospheric pressure by applying a high voltage (800 –1000 V) to a platinum (Pt) wire inserted into the solution. Small droplets and ions produced by nanoES are sampled into the mass spectrometer through a heated metal capillary. Gaseous ions are transmitted through the skimmer and accumulated in hexapole for a certain time to enhance signal-to-noise (S/N) ratio. After accumulation, ions are ejected and accelerated by a high voltage through the fringing field of a 4.7 T superconducting magnet. Then the ions are decelerated and eventually captured by a combination of electric and magentic field in FT-ICR cell for detection. The typical base pressure for the instrument is ~ 5 x 10^{-10} mbar, maintained by the differential pumping system.

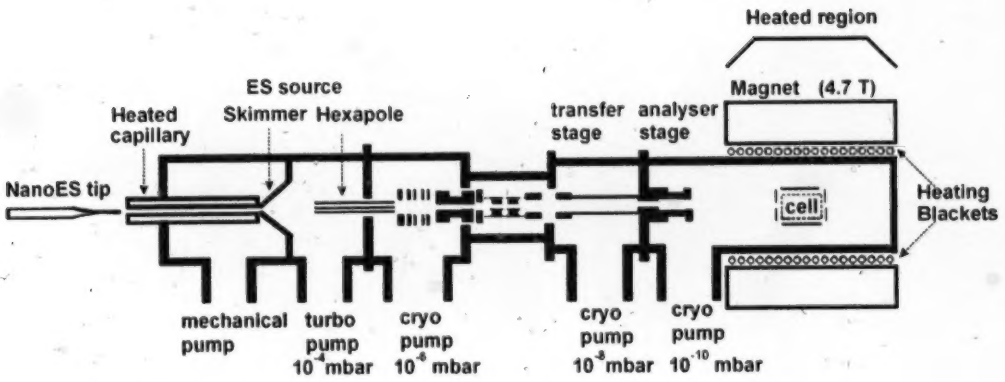


Figure 1.4 NanoES-FT-ICR/MS instrument diagram.

To perform time-resolved experiments on the ions of interest, the ions are isolated using a single rf frequency and broadband rf sweep excitation. Then, the isolated ions are stored for variable reaction times prior to detection. During the reaction time, ions undergo dissociation by a variety of dissociation techniques such as multiple-collision activation dissociation, infrared multiphoton dissociation, electron capture dissociation and BIRD that was used in the present work. The principle of BIRD will be described in detail in the following section.

1.2.2 Blackbody Infrared Radiative Dissociation (BIRD)

BIRD is a dissociation technique which allows ions to undergo unimolecular dissociation by blackbody radiation at essentially zero pressure. Since the ultra low pressure is maintained in the FT-ICR cell, BIRD is implemented with FT-ICR.

The BIRD mechanism was originally proposed by Perrin in 1919 [41] to explain unimolecular dissociation processes. The hypothetical molecular ion, AB^+, is used to

describe the radiative dissociation mechanism. As shown in eq. 1.2a-c [42], the unimolecular reaction of AB^+ producing A^+ and B (eq. 1.2a) involves two steps. AB^+ is excited to an activated state AB^{+*} by the absorption of infrared photons, eq.1.2b. Then, the AB^{+*} could relax back into AB^+ through the emission of infrared photons, or dissociate into the products A^+ and B (eq.1.2c). Under steady state conditions, combining the three equations, the observed rate constant of unimolecular reaction, k_{uni} is expressed in terms of the radiative rate constants, $k_{1,rad}$ and $k_{-1,rad}$, and the unimolecular dissociation rate constant, k_d (eq.1.3a). If k_d is much larger than $k_{-1,rad}$, k_{uni} is only related to $k_{1,rad}$ and $k_{-1,rad}$ (eq.1.3b). Therefore, the observed reaction kinetics reflect only the rate of photon absorption, and no information about the dynamics of the dissociation process. If k_d is much smaller than $k_{-1,rad}$, k_{uni} is determined by the dissociation kinetics of ions with a Boltzmann distribution of internal energies (eq. 1.3c) [43], and the Arrehnius activation energy (E_a) and preexponential factor (A) can be determined from the temperature dependence of k_{uni}. Large ions such as protein ions are in the rapid-exchange limit (REX), wherein the energy exchange is rapid compared to the dissociation kinetics and the internal energies of ions are characterized by a Boltzmann distribution.

$$AB^+ \xrightarrow{k_{uni}} A^+ + B \qquad (1.2a)$$

$$AB^+ + h\nu \underset{k_{-1,rad}}{\overset{k_{1,rad}}{\rightleftarrows}} AB^{+*} \qquad (1.2b)$$

$$AB^{+*} \xrightarrow{k_d} A^+ + B \qquad (1.2c)$$

$$k_{uni} = k_d\left(\frac{k_{1,rad}}{k_{-1,rad} + k_d}\right) \qquad (1.3a)$$

$$k_{uni} = k_{1,rad}, \text{ if } k_d >> k_{-1,rad} \quad (1.3b)$$

$$k_{uni} = k_d \frac{k_{1,rad}}{k_{-1,rad}}, \text{ if } k_d << k_{-1,rad} \quad (1.3c)$$

Two essential conditions are required for BIRD to be experimentally observable; first, the pressure of ICR cell is low enough (below 10^{-6} Torr) so that collisional energy exchange between ions and surrounding is negligible; and second, that the time scale of observation of the dissociating molecules is long enough (on the order of seconds) for this relatively slow mechanism to give a significant degree of dissociation of the population [44].

BIRD, as a valuable method for studying thermal kinetics in the gas phase, has been successfully applied to relatively small biomolecules and more recently to oligonucleotide complexes, protein–ligand complexes and protein-protein complexes [43]. It has been demonstrated that the BIRD method combined with functional group replacement can be used to identify and quantify the individual interactions in the gaseous biomolecular complexes [30e-h].

1.3 The Present Work

Although many studies of using nanoES-MS to measure binding affinities of biomolecular complexes and to study specificity of noncovalent complexes have been demonstrated, this approach is yet to be widely adopted. The principal perceived weakness is that the MS data do not reflect the equilibrium composition in solution under certain experimental conditions. To solve this problem, it is necessary to understand the underlying causes of the problem. Therefore, the present work is mainly focused on the

influence of solution and the gas phase processes on the determination of protein-ligand binding affinities using nanoES-FT-ICR/MS.

In Chapter 2, an extensively studied protein-carbohydrate complex (a single chain variable fragment of a monoclonal antibody (scFv) and its native trisaccharide ligand) was chosen as a model system to examine the influence of solution and gas phase processes on the measurement of protein-carbohydrate association constants using nanoES-FT-ICR/MS. Several factors, such as changes in solution pH and analyte concentration during the nanoES process and in-source dissociation of the gaseous complexes were systematically investigated. Experimental conditions were identified that preserve the solution-phase distribution of bound to unbound protein in the gas phase, such that the nanoES mass spectrum provides a quantitative measure of the solution-phase binding equilibrium composition. These conditions include the use of short spray durations (< 10 min) to minimize pH changes, equimolar concentrations of protein and ligand to minimize the formation of nonspecific complexes, and short accumulation times (< 2 s) in the hexapole of the ion source to avoid collisional heating and dissociation of the gaseous complex. Application of this methodology to the scFv and a series of carbohydrate ligands yields results that are in agreement with values previously determined by isothermal titration calorimetry. Competitive binding experiments performed on solutions containing the scFv and a mixture of carbohydrate ligands were also found to yield accurate association constants.

During the course of the aforementioned study, we found that nonspecific protein-carbohydrate complexes are readily formed by nanoES, particularly at high carbohydrate concentrations. The formation of nonspecific complexes can complicate ES-MS

experiments in a number of ways, such as obscuring the binding stoichiometry of specific complexes, and leading to higher binding affinities measured by ES-MS. To answer the question whether nonspecific complexes can be selectively removed in the gas phase, a comparative study of the kinetic and energetic stability of specific and nonspecific complexes in the gas phase was undertaken. In Chapter 3, nanoES was performed on aqueous solutions containing the scFv and its native ligand at a high concentration. The protonated ions produced corresponded to scFv complexed with one and two ligands. Since the scFv is known to have a single specific binding site for its native ligand, the 1:2 complex ions must originate from nonspecific binding during the nanoES process. Time-resolved thermal dissociation experiments were performed on the protonated 1:1 and 1:2 complex ions with +10 and +11 charges, produced from solutions with a high carbohydrate concentration. BIRD of these ions proceeds exclusively by the loss of neutral carbohydrate. Nonlinear, first order dissociation kinetics were obtained for the 1:1 ions, suggesting that there are two structurally distinct complex ions in the gas phase. The dissociation of the 1:2 complex ions occurs via parallel pathways involving the loss of a specific or nonspecific carbohydrate ligand. The measured rate constant from the linear kinetics at a given temperature corresponds to the sum of rate constants for the loss of the specific and nonspecific ligand. Comparison of these kinetic data with results obtained for 1:1 ions produced at low ligand concentration [30e] indicates that complexes arising from nonspecific binding exhibit greater kinetic stability than the complex produced from dilute solution. Furthermore, the E_a determined for the nonspecific ions is greater than the value reported for the corresponding specific complex at +10 charge state, and approximately the same for the +11 charge state. These results indicate that nonspecific

binding can lead to complexes that are both kinetically and energetically more stable than complexes which originated from specific interactions in solution. This study revealed that it may not be feasible to selectively remove the nonspecific protein-ligand complexes in the gas phase.

In order to understand the mechanism governing the formation of nonspecific protein-carbohydrate complexes, factors influencing the formation of nonspecific protein-carbohydrate complexes during nanoES have been investigated in Chapter 4. Protonated and deprotonated nonspecific complexes of ubiquitin (Ubq) and protonated complexes of bovine carbonic anhydrase II (CA) with carbohydrates, ranging in size from mono- to tetrasaccharide, were produced by nanoES and detected with a FT-ICR/MS. Both the fraction of protein engaged in nonspecific binding with the carbohydrates and the number of carbohydrates bound to the protein increase with increasing carbohydrate concentration. At a given concentration of protein and carbohydrate, nonspecific binding is favoured for small (mono- and disaccharide) or hydrophilic carbohydrates over larger or more hydrophobic molecules, which tend to form gaseous monomer or cluster ions by nanoES. However, the extent of nonspecific binding is insensitive to the structure of the protein, with similar distributions of nonspecific complexes observed for both CA and Ubq. Nonspecific association is also insensitive to the charge state of the complex. A comparable degree of binding is observed for complexes in their protonated and deprotonated forms. Furthermore, the number of bound ligands can exceed significantly the charge state of the complex. Thermal dissociation experiments performed on the gaseous nonspecific complexes reveal that their kinetic stability is sensitive to both the structure of the carbohydrate (*i.e.* mono < di < tri < tetrasaccharide) and the protein (Ubq

< CA) and to the charge state, although no simple relationship between stability and charge state was identified. Taken together, the results of this study suggest that neutral protein-carbohydrate interactions (*e.g.* hydrogen bonds) contribute significantly and, perhaps, predominantly to the formation and stabilization of the nonspecific complexes. A strategy to minimize the formation of the nonspecific complexes, which is based on the enhancement of gaseous carbohydrate ion formation through the addition of metal salts (*e.g.* $CaCl_2$) to the nanoES solution, is demonstrated.

To further investigate the nature of nonspecific binding in the gas phase, BIRD experiments have been performed on a series of protonated and deprotonated 1:1 and protonated 1:2 protein-carbohydrate complexes formed by nonspecific interactions during the nanoES process. Nonspecific interactions between the proteins CA, Ubq and bovine pancrease trypsin inhibitor (BPTI) and several carbohydrates, ranging in size from mono- to tetrasaccharide, have been investigated. Over the range of temperatures studied (60 – 190 °C), BIRD of the protonated and deprotonated complexes proceeds exclusively by the loss of the carbohydrate in its neutral form. The rates of dissociation of the 1:1 complexes containing a mono- or disaccharide decrease with reaction time, suggesting the presence of two or more kinetically distinct structures, produced either during the nanoES or dissociation processes. In contrast, the 1:1 complexes of the tri- and tetrasaccharides exhibit simple first-order dissociation kinetics, a result that, on its own, is suggestive of a single preferred carbohydrate binding site or multiple equivalent sites in the gas phase. A comparative analysis of dissociation kinetics measured for protonated 1:1 and 1:2 complexes of Ubq with αTal[αAbe]αMan further supports the presence of a single preferred binding site. However, a similar analysis performed on the complexes of

CA and αTal[αAbe]αMan suggests equivalent but dependent carbohydrate binding sites exist in the gas phase. Analysis of the Arrhenius parameters (E_a and A) determined for the dissociation of 1:1 complexes of CA with structurally-related trisaccharides provides compelling evidence that neutral intermolecular hydrogen bonds contribute, at least in part, to the stability of the gaseous complexes. Surprisingly, the energetic stability of the complexes of the same charge state is not strongly sensitive to the structure (primary or higher order) of the protein, suggesting that the carbohydrates are able to form energetically equivalent interactions with the various functional groups presented by the protein. For a given protein-carbohydrate complex, the energetic stability is sensitive to charge state, although no simple relationship between E_a and charge state is evident. It is proposed that both ionic and neutral hydrogen bonds are responsible for stabilizing nonspecific protein-carbohydrate complexes in the gas phase and that the relative contribution of the neutral and ionic interactions is strongly influenced by charge state with neutral interactions dominating at low charge states and ionic interactions dominating at high charge states.

1.4 Literature Cited

(1) (a) Cleveland, D. W.; Mao, Y.; Sullivan, K. F. *Cell* **2003**, *112*, 407-421; (b) Helnius, A.; Aebi, M. *Science* **2001**, *23*, 2364-2369; (c) Igakura, T.; Stinchcombe, J. C.; Goon, P. K. C.; Taylor, G. P.; Weber, J. N.; Griffiths, G. M.; Tanaka, Y.; Osame, M.; Bangham, C. R. M. *Science* **2003**, *299*, 1713-1716; (d) Huang, J. F.; Yang, Y.; Sepulveda, H.; Shi, W. X.; Hwang, I.; Peterson, P. A.; Jackson, M. R.; Sprent, J.; Cai, Z. L. *Science* **1999**, *286*, 952-954.

(2) (a) Su, T. M.; Yang, Y. S. *Biochemistry* **2003**, *42*, 6863-6870; (b) Pletneva, E. V.; Laederach, A. T.; Fulton, D. B.; Kostic, N. M. *J. Am. Chem. Soc.* **2001**, *26*, 6232-6245; (c) Worrall, J. A. R.; Kolczak, U.; Canters, G. W.; Ubbink, M. *Biochemistry* **2001**, *24*, 7069-7076; (d) Zuiderweg, E. R. P. *Biochemistry* **2002**, *41*, 1-7; (e) Pickford, A. R.; Campbell, I. D. *Chem. Rev.* **2004**, *104*, 3557-3565; (f) Coille, I.; Gauglitz, G.; Hoebeke, J. *Anal. Bioanal. Chem.* **2002**, *372*, 293-300; (g) Yang, H. W.; Xiao, Y.; Lu. Y.; Chen, Y. H. *Immonobiology* **2001**, *203*, 778-785; (h) Madrid, K. P.; De Crescenzo, G.; Wang, S. W.; Jardim, A. *Mol. Cell. Biol.* **2004**, *24*, 7331-7344; (i) Silvaggi, N. R.; Kaur, K.; Adediran, S. A.; Pratt, R. F.; Kelly, J. A. *Biochemistry* **2004**, *43*, 7046-7053.

(3) McLafferty, F. W. *Science* **1981**, *214*, 280-287.

(4) (a) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71; (b) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, *88*, 4451-4459; (c) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, *88*, 4671-4675.

(5) (a) Wilm, M. S.; Mann, M. *Int. J. Mass Spectrom. Ion Processes* **1994**, *136*, 167-180; (b) Wilm, M. S.; Mann, M. *Anal. Chem.* **1996**, *68*, 1-8.

(6) Ganem, B.; Li, Y.-T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294-6296.

(7) Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1991**, *113*, 8534-8535.

(8) Ganem, B.; Li, Y.-T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 7818-7819.

(9) Baca, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1992**, *114*, 3992-3993.

(10) Drummond, J. T.; Ogorzalek Loo, R. R.; Matthews, R. G. *Biochemistry* **1993**, *32*, 9282-9289.

(11) Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605-611.

(12) Light-Wahl, K. J.; Schwartz, B. L.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 5271-5278.

(13) Fizgerald, M. C.; Chernushevich, I.; Standing, K. G.; Kent, S. B. H.; Whitman, C. P. *J. Am. Chem. Soc.* **1995**, *117*, 11075-11080.

(14) Ganguly, A. K.; Pramanik, B. N.; Tsarbopoulos, A.; Covey, T. R.; Huang, E.; Fuhrman, S. A. *J. Am. Chem. Soc.* **1992**, *114*, 6559-6560.

(15) Rostom, A. A.; Robinson, C. V. *J. Am. Chem. Soc.* **1999**, *121*, 4718-4719.

(16) Griffith, M. C.; Risen, L. M.; Greig, M. J.; Lesnik, E. A.; Sprankle, K. G.; Griffey, R. H.; Freier, S. M. Rostom, A. A.; Robinson, C. V. *J. Am. Chem. Soc.* **1995**, *117*, 831-832.

(17) Greig, M. J.; Gaus, H.; Cummins, L. L.; Sasmor, H.; Griffey, R. H. *J. Am. Chem. Soc.* **1995**, *117*, 10765-10766.

(18) Light-Wahl, K. J.; Springer, D. L.; Winger, B. E.; Edmonds, C. G.; Camp, D. G.; Thrall, B. D.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, *115*, 803-804.

(19) Light-Wahl, K. J.; Winger, B. E.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, *115*, 5869-5870.

(20) Ganem, B.; Li, Y. T.; Henion, J. D. *Tetrahedron lett.* **1993**, *34*, 1445-1448.

(21) Gale, D. C.; Goodlett, D. R.; Light-Wahl, K. J.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 6027-6028.

(22) Tito, M. A.; Tars, K.; Valegard, K.; Hajdu, J.; Robinson, C. V. *J. Am. Chem. Soc.* **2000**, *122*, 3550-3551.

(23) (a) Smith, R. D., Light-Wahl, K. J. *Biol. Mass Spectrom.* **1993**, *22*, 493-501; (b) Smith, R. D., Zhang, Z. *Mass Spectrom. Rev.* **1994**, *13*, 411-429; (c) Loo, J. A. *Bioconj. Chem.* **1995**, *6*,644-665; (d) Przybylski, M.; Glocker, M. O. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 807-826; (e) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(24) (a) Cheng, X. H.; Harms, A. C.; Goudreau, P. N.; Terwilliger, T. C.; Smith, R. D. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 7022-7027; (b) Feng, R.; Castelhano, A. L.; Billedeau, R.; Yuan, Z. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1105-1111; (c) Ganem, B.; Li, Y. T.; Hsieh, Y. L.; Henion, J. D. Kaboord, B. F.; Frey, M. W.; Benkovic, S. J. *J. Am. Chem. Soc.* **1994**, *116*, 1352-1358.

(25) (a) Cheng, X.; Chen, R.; Bruce, J. E.; Schwartz, B. L.; Anderson, G. A.; Hofstadler, S. A.; Gale, D. C.; Smith, R. D.; Gao, J.; Sigal, G. B.; Mammen, M.; Whitesides, G. M. *J. Am. Chem. Soc.* **1995**, *117*, 8859-8860; (b) Loo, J. A.; Holler, T. P.; Foltin, S. K.; McConnell, P.; Banotai, C. A.; Horne, N. M.; Mueller, W. T.; Stevenson, T. I.; Mack, D. P. *Proteins: Struct. Funct. Genet.* Suppl **1998**, *2*, 28-37; (c) Rostom, A. A.; Tame, J. R. H.; Ladbury, J. E.; Robinson. C. V. *J. Mol. Biol.* **2000**, *296*, 269-279; (d) Hill, T. J.; Lafitte, D.; Wallace, J. I.; Cooper, H. J.; Tsvetkov, P. O.; Derrick, P. J. *Biochemistry* **2000**, *39*, 7284-7290.

(26) (a) Greig, M. J.; Gaus, H; Cummins, L. L; Sasmor, H.; Griffey, R. H. *J. Am. Chem. Soc.* **1995**, *117*, 10765-10766; (b) Loo, J. A.; Hu, P.; McConnell, P.; Mueller, W. T.; Sawyer, T. K.; Thanabal, V. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 234-243; (c) Ayed, A; Krutchinsky, A. N.; Ens, W.; Standing, K. G.; Duckworth, H. W. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 339-344; (d) Sannes-Lowery,

K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264-271; (e) Kitova, E. N.; Kitova, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605-611; (f) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945-4955; (g) Zhang, S.; Van Pelt, C. K.; Wilson, D. B. *Anal. Chem.* **2003**, *75*, 3010-3018; (h) Hagan, N.; Fabris, D. *Biochemistry*, **2003**, 42, 10736-10745; (i) Tjernberg, A.; Carnö, S.; Oliv, F.; Benkestock, K.; Edlund, P.-O.; Griffiths, W. J.; Hallén, D. *Anal. Chem.* **2004**, *76*, 4325-4331.

(27) (a) Daneshfar, R; Kitova, E. N.; Klassen, J. S. *J. Am. Chem. Soc.* **2004**, *126*, 4786-4787; (b) Benesch, J. L. P.; Sobott, F.; Robinson, C. V. *Anal. Chem.* **2003**, *75*, 2208-2214.

(28) (a) Gao, J.; Cheng, X.; Chen, R.; Sigal, G. B.; Bruce, J. E.; Schwartz, B. L.; Hofstadler, S. A.; Anderson, G. A.; Smith, R. D.; Whitesides, G. M. *J. Med. Chem.* **1996**, *39*, 1949-1955; (b) Wigger, M.; Eyler, J. R.; Benner, S. A.; Li, W. Q.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1162-1169.

(29) (a) McCammon, M. G.; Scott, D. J.; Keetch, C. A.; Greene, L. H.; Purkey, H. E.; Petrassi, H. M.; Kelly, J. W.; Robinson, C. V. *Structure* **2002**, *10*, 851-863; (b) Aquilina, J. A.; Benesch, J. L. P.; Bateman, O. A.; Slingsby, C.; Robinson, C. V. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10611-10616; (c) Sobott, F.; Robinson, C. V. *Int. J. Mass Spectrom.* **2004**, *236*, 25-32.

(30) (a) Hunter, C. L.; Mauk, A. G.; Douglas, D. J. *Biochemistry* **1997**, *36*, 1018-1025; (b) Gabelica V.; De Pauw, E. *J. Mass Spectrom.* **2001**, *36*, 397-402; (c) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, *120*, 9605-9613; (d) Felitsyn, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2001**,

73, 4647-4661; (e) Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 5902-5913; (f) Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 9340-9341; (g) Kitova, E. N.; Wang, W.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 13980-13981; (h) Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *Angew. Chem. Int. Ed.* **2004**, *43*, 4183-4186; (i) Mauk; M. R.; Mauk, A. G.; Chen, Y.-L.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 59-71; (j) Catalina M. I.; de Mol N. J.; Fischer M. J. E.; Heck A. J. R. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2572-2579.

(31) (a) Light-Wahl, K. J.; Schwartz, B. L.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 5271-5278; (b) Nettleton, E. J.; Sunde, M.; Lai, Z.; Kelly, J. W.; Dobson, C. M.; Robinson, C. V. *J. Mol. Biol.* **1998**, *281*, 553-564; (c) Rostom, A. A.; Robinson, C. V. *J. Am. Chem. Soc.* **1999**, *121*, 4718-4719; (d) McCammon, M. G.; Hernández, H.; Sobott, F.; Robinson, C. V. *J. Am. Chem. Soc.* **2004**, *126*, 5950-5951.

(32) Daniel. J. M.; Firess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1-27.

(33) (a) Cunniff, J. B.; Vouros, P. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 437-447; (b) Robison, C. V., Chung, E. W., Kragelund, B. B., Knudsen, J., Aplin, R. T., Poulsen, F. M.; Dobson C. M. *J. Am. Chem. Soc.* **1996**, *118*, 8646-8653; (c) Huang, H.-W.; Wang, K.-T. *Biochem. Biophys. Res. Commun.* **1996**, *227*, 615-621; (d) Wu, Q. Y.; Gao, J. M.; JosephMcCarthy, D., Sigal, G. B., Bruce, J. E., Whitesides, G. M., Smith, R. D. *J. Am. Chem. Soc.* **1997**, *119*, 1157-1158; (e) Nesatyy, V. J. *J. Mass Spectrom.* **2001**, *36*, 950-959.

(34) Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, *65*, 972A-986A.

(35) (a) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, 64, 2287-2294; (b) Thomson, B. A.; Iribarne, J. V. *J. Chem. Phys.* **1979**, 71, 4451-4463.

(36) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240-2249.

(37) (a) de la Mora, J. F.; Van Berkel; G. J.; Enke, C. G.; Cole, R. B.; Martinez-Sanchez, M.; Fenn, J. B. *J. Mass Spectrom.* **2000**, *35*, 939-952; (b) Van Berkel, G. J.; Asano, K. G.; Schnier, P. D. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 853-862.

(38) Wang, H.; Agnes, G. R. *Anal. Chem.* **1999**, *71*, 3785-3792.

(39) Lim, H.–K.; Hsieh, Y.; Ganem, B.; Henion, J. *J. Mass Spectrom.* **1995**, *30*, 708-714.

(40) (a) Marshall, A. G.; Grosshans, P. B. *Anal. Chem.* **1991**, *63*, 215A-229A; (b) Buchanan, M. V.; Hettich, R. L. *Anal. Chem.* **1993**, *65*, 245A-259A; (c) Jonathan Amster, I. *J. Mass Spectrom.* **1996**, *31*, 1325-1337; (d) Marshall, A. G.; Hendrickson, C. L. *Int. J. Mass Spectrom.* **2002**, *215*, 59-75.

(41) Perrin, J. *Ann. Phys.* **1919**, *11*, 5-108.

(42) Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644.

(43) Price, W. D.; Williams, E. R. *J. Phys. Chem. A* **1997**, *101*, 8844-8852.

(44) Dunbar, R. C. *Mass Spectrom. Rev.* **2004**, *23*, 127-158.

Chapter 2

Influence of Solution and Gas Phase Processes on Protein-Carbohydrate Binding Affinities Determined by Nanoelectrospray Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry*

2.1 Introduction

The specific recognition of carbohydrates by proteins underlies many important biological processes such as cellular growth and adhesion, bacterial and viral infections, inflammation and the immune response [1-3]. To understand the nature of the protein-carbohydrate recognition processes, detailed information regarding the specificity and affinity of the interactions is required. There are several quantitative techniques that can be used to evaluate protein-carbohydrate binding, including ITC [4], SPR [5] and frontal affinity chromatography–mass spectrometry (FAC/MS) [6]. Each of these techniques has particular advantages and disadvantages. ITC is probably the most widely used method for determining protein-carbohydrate binding affinities and is the only technique that provides a direct measure of the enthalpy of association. However, ITC has several limitations, including the often-prohibitive requirement for milligram quantities of both protein and ligand for each analysis and the inability to provide direct information on binding stoichiometry or to distinguish ligand binding to different protein quaternary structures. SPR offers good sensitivity and can be used to evaluate the rate constants for the association-dissociation reactions. FAC/MS is also quite sensitive and can be used to rapidly determine binding constants for carbohydrate ligands present as mixtures. A

A version of this chapter has been published: Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945-4955.

common disadvantage of both the SPR and FAC/MS techniques is the requirement for the immobilization of one of the analytes (protein or ligand) on a solid surface. In addition to being impractical in certain cases, immobilization may alter the nature of the protein-carbohydrate interactions leading to incorrect binding affinities.

MS combined with a soft ionization technique, with its speed, sensitivity, specificity and ability to directly determine binding stoichiometry has emerged as a powerful tool for studying non-covalent biomolecular complexes, including protein assemblies and protein-ligand complexes [7]. Of the available ionization techniques, ES and nanoES have proven to be the most useful for investigating interactions between soluble, biologically-relevant molecules [8]. NanoES, which operates at lower solution flow rates than ES and emits correspondingly smaller droplets (estimated to be approximately 100-200 nm in diameter) [9] has been shown to be a particularly powerful technique for investigating biomolecular complexes. In addition to being more sensitive than ES, normally requiring picomoles or less of analyte per analysis, nanoES readily allows for the transfer of non-covalent complexes from buffered aqueous solution to the gas phase and, therefore, can be used to study binding under near-physiological conditions. Furthermore, the short lifetimes of the nanoES droplets, which are estimated to be tens of μs, is likely advantageous for preserving the original solution composition throughout the formation of gaseous ions, *vide infra*.

The first direct observation of a protein-carbohydrate complex, hen egg white lysozyme and a hexasaccharide of N-acetylglucosamine, by ES-MS was reported in 1991 [10]. Heck and coworkers used ES-MS to study the binding of carbohydrate ligands with the lectin, apo-concanavalin (Con A) [11] which exists in both dimer and tetramer forms

in solution. From the ES-MS data, the authors evaluated the binding stoichiometry for the different quaternary Con A complexes and demonstrated that the dimer and tetramer forms exhibit similar affinities for the carbohydrate ligands investigated. In a recent study by Siebert *et al.*, the binding of 15-mer peptides (*i.e.* synthetic mini-proteins) with disaccharide ligands was investigated by nanoES-MS [12]. From the ES-MS data, the binding affinities of the ligands were estimated to be on the order of $1 \times 10^3\ M^{-1}$.

ES- and nanoES-MS also have tremendous potential as a rapid and sensitive tool for quantifying binding affinities for biomolecular complexes. A number of quantitative studies have appeared in recent years dealing with protein-protein [13] and protein-peptide complexes [14], protein-oligonucleotide complexes [15], and peptide and RNA-binding antibiotics [16] and small molecule-RNA complexes [17]. Zenobi and coworkers have described these studies in a recent review article [7]. The first quantitative study of protein-carbohydrate binding involved the interaction between the P^k trisaccharide and the B_5 homopentamer of the Shiga-like toxin I [18].

While the success of the aforementioned studies indicates that ES- and nanoES-MS can serve as a general tool for quantifying non-covalent interactions, this technique has yet to be widely adopted. The principal perceived weakness of the approach is the sensitivity of the relative abundance of gas phase ions (*i.e.* bound and unbound forms) observed in the mass spectrum, and hence the binding constants, to the choice of experimental conditions. The ratio of complex to protein may be influenced by solution phase processes, such as an increase in analyte concentration in the droplets resulting from solvent evaporation, as well as other changes in solution composition due to electrochemical reactions [19, 20]. The formation of nonspecific complexes during the

nanoES process [21] and dissociation of the gaseous complex in the ion source [22] can also influence the distribution of ions in the mass spectrum. Consequently, *de novo* affinity measurements normally require the use of a reference complex with a known binding constant to serve as an internal standard or for "tuning" the experimental conditions to yield meaningful values.

In the case of protein-carbohydrate binding, there are additional factors that may hinder the use of ES-MS for detection and quantification. One factor is the low binding constants characteristic of protein-carbohydrate complexes, generally in the range of 10^3 to 10^5 M^{-1} and even lower for some monosaccharide ligands [23]. With few exceptions, ES- and nanoES-MS studies of protein-ligand complexes have been restricted to moderately or strongly bound complexes, with association constants greater than 10^5 M^{-1}. The detection of low affinity complexes requires the use of high analyte (usually ligand) concentrations, which may suppress the formation of gaseous protein and protein-ligand ions and enhance the formation of nonspecific complexes [21], which obscures the true binding affinity and stoichiometry. Heterogeneous protein structure and composition, characteristic of lectins (carbohydrate-binding proteins) [1], results in a distribution of ions with similar *m/z* and may further complicate analysis. This, combined with the low molecular weight of many model carbohydrate ligands (mono- to tetrasaccharides), may require the use of mass analyzers with relatively high-resolution capabilities.

The present work describes the influence of gas phase and solution processes on protein-carbohydrate binding affinity measurements performed with nanoES and FT-ICR/MS. A genetically engineered scFv, based on the carbohydrate-binding antibody Se155-4 [24] and its native trisaccharide ligand, α-D-Gal(1→2)[α-D-Abe(1→3)α-D-

Man*p*→OMe ≡ αGal[αAbe]αMan, served as a model system for this study. The complexes of Galα[Abe]Man with the Se155-4 antibody (IgG), the antigen-binding fragment (Fab) and the scFv have been the focus of extensive investigation in solution as part of a comprehensive study of carbohydrate recognition by proteins. Due to molecular weight considerations, the scFv•αGal[αAbe]αMan complex was chosen for this study instead of the corresponding complex of the IgG and Fab proteins. Association constants (K_{assoc}) for the binding of the scFv, Fab and IgG with the native trisaccharide antigen have been previously determined at 298 K by ITC and the values range from 1.1 to 2.1 x 10^5 M^{-1} [24, 25]. The most recent value, determined for the IgG, is (1.6 ± 0.2) x 10^5 M^{-1} [26]. The goal of the present work was to identify experimental conditions that preserve the equilibrium distribution of bound (scFv•αGal[αAbe]αMan) and unbound protein (scFv) in solution, into the gas phase, such that the nanoES-mass spectrum provides a quantitative measure of the solution composition and K_{assoc}. To demonstrate that experimental conditions were generally suitable for protein-carbohydrate complexes, K_{assoc} for the scFv and three other oligosaccharide ligands, structural analogs of the native ligand, have been measured with this approach and compared with values obtained by ITC. These ITC-derived K_{assoc} values were determined for binding with the IgG, but not the scFv, protein [26, 27]. However, because of the uniform binding mode for the complexes of the IgG, Fab and scFv proteins, the binding constants determined for the IgG complexes should closely resemble those of the corresponding scFv complexes. The application of competitive binding experiments, wherein binding affinities of several carbohydrate ligands are determined simultaneously, is also described.

2.2 Experimental

2.2.1 Protein and Carbohydrate Ligands

The carbohydrate-binding antibody single chain fragment, scFv (MW 26539 Da), was produced using recombinant technology [24]. The scFv was concentrated and dialyzed against deionized water using MICROSEP microconcentrators with a molecular weight cut-off of 10 kDa, and lyophilized prior to MS analysis. The scFv was weighed immediately after removing it from the lyophilizer, dissolved in a known volume of aqueous 50 mM ammonium acetate and stored at –20 °C if not used immediately. The structures of the ligands are shown in Figure 2.1. Any adsorbed water was removed from the ligands prior to the preparation of stock solutions by drying the ligands in a vacuum chamber maintained at ~5 torr and 56 °C. Each nanoES solution was prepared from stock solutions of scFv and ligand with known concentrations. A 50 mM aqueous solution of ammonium acetate was added to yield a final concentration of 1 mM. The pH of the nanoES solution was determined using an ORION 710Aplus pH/ISE meter with microelectrodes.

2.2.2 Mass Spectrometry

All experiments were performed using an Apex II 47e FT-ICR mass spectrometer (Bruker, Billerica, MA) equipped with a modified external nanoES source. This instrument has been described previously [28] and only a brief description is given here. NanoES tips were pulled from aluminosilicate tubes (1 mm o.d., 0.68 mm i.d.) using a P-

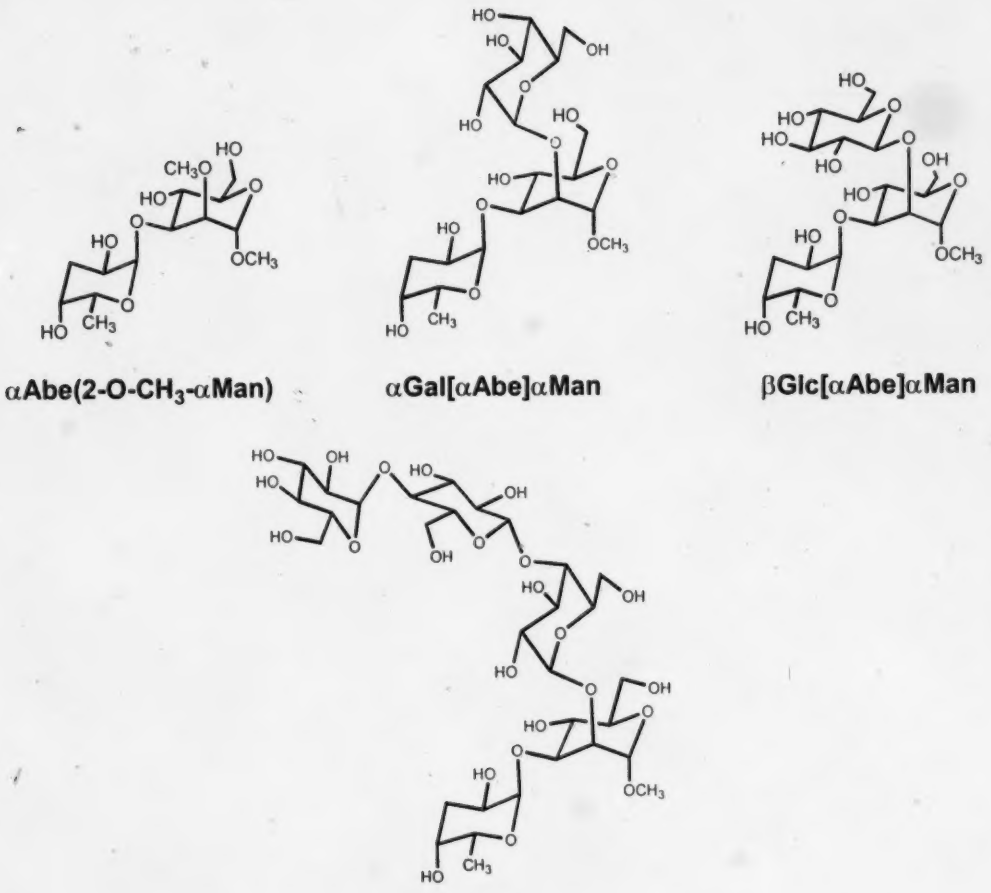


Figure 2.1 Structures of the oligosaccharide ligands.

2000 or P-97 micropipette puller (Sutter Instruments, Novato, CA). Shown in Figure 2.2 are images, acquired with a scanning electron microscope (SEM), of two typical nanoES tips pulled under identical conditions. The shape of the tips was found to be quite reproducible, with an outer diameter of 4 to 7 μm. The wall thickness at the end of the tips was determined to be approximately 75 nm such that the i.d. and o.d. at the end of the

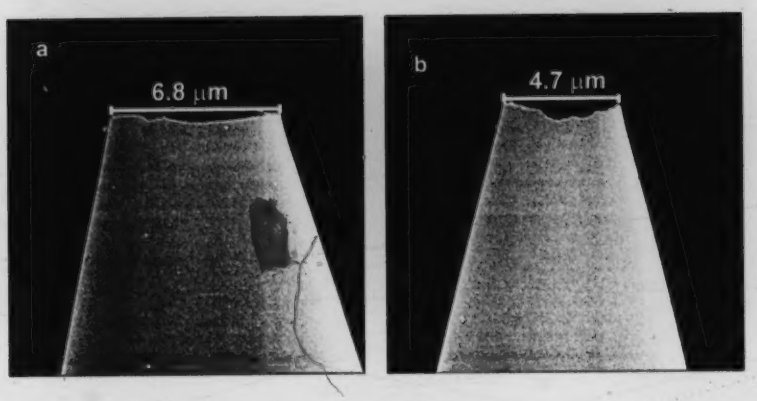


Figure 2.2 SEM images of two nanoES tips.

nanoES tips are similar. A Pt wire, inserted into the other end of the nanoES tip, was used to establish electrical contact with the nanoES solution. A potential of 800 to 1100 V was applied to the Pt wire in the nanoES tip in order to spray the solution. The tip was positioned 1-2 mm from a stainless steel sampling capillary using a microelectrode holder. The solution flow rate ranged from 20 to 50 nL/min, depending on the outer diameter of the nanoES tip and the voltage. Typically, a stable nanoES ion current of ~0.1 μA was achieved.

The droplets and gaseous ions produced by nanoES were introduced into the vacuum chamber of the mass spectrometer through a heated stainless steel sampling capillary (0.43 mm i.d.) maintained at an external temperature of 66 °C. The gaseous ions sampled by the capillary (52 V) were transmitted through a skimmer (4 V) and accumulated in a hexapole (600 V p-p). Unless otherwise noted, an accumulation time of 1.5 s was used. The ions were subsequently ejected from the hexapole and injected at ~ -2700 V into the bore of a 4.7 tesla superconducting magnet, decelerated and introduced into the ion cell. The typical base pressure for the instrument was ~5 x 10^{-10} mbar. Data acquisition was performed using the Bruker Daltonics XMASS software (version 5.0).

The time-domain spectra consisted of the sum of 30 transients containing 128 K data points per transient.

2.3 Results and Discussion

2.3.1 Determining K_{assoc} by MS

The general equilibrium expression for the association reaction involving a protein (P) and ligand (L) (eq. 2.1) is given by eq. 2.2:

$$P + L \leftrightarrow P{\bullet}L \tag{2.1}$$

$$K_{assoc} = [PL]_{equil} / [P]_{equil}[L]_{equil} \tag{2.2}$$

The equilibrium concentrations, $[PL]_{equil}$, $[P]_{equil}$ and $[L]_{equil}$, can be deduced from the initial concentration of protein and ligand in solution, $[P]_o$ and $[L]_o$, and the relative abundance of the bound and unbound protein ions, $P{\bullet}L^{n+}$ and P^{n+}, measured in the mass spectrum. Assuming that the spray and detection efficiencies for the $P{\bullet}L^{n+}$ and P^{n+} ions are similar, the ratio (R) of the ion intensity (I) of the bound and unbound protein ions determined from the mass spectrum should be equivalent to the ratio of the concentrations in solution at equilibrium as shown in eq. 2.3:

$$R = I(P{\bullet}L^{n+})/I(P^{n+})) = [PL]_{equil}/[P]_{equil} \tag{2.3}$$

The spray efficiencies and detection efficiencies are expected to be similar when the solvent exposed surface (area and properties) and the molecular weight of the protein and complex are comparable. These conditions apply to the scFv•αGal[αAbe]αMan system, since the ligand interacts with only 7% of the protein surface (estimated from crystal structure) and the difference in molecular molecular weight is less than 2%.

The equilibrium concentration, $[PL]_{equil}$, can be determined from the value of R and $[P]_o$ using the following expression:

$$[PL]_{equil} = \frac{R[P]_o}{1+R} \tag{2.4}$$

The equilibrium concentration $[L]_{equil}$ can be found from eq. 2.5 and K_{assoc} can then be determined with eq. 2.6.

$$[L]_{equil} = [L]_o - [PL]_{equil} \tag{2.5}$$

$$K_{assoc} = \frac{[PL]_{equil}}{[P]_{equil}([L]_o - [PL]_{equil})} \tag{2.6a}$$

$$K_{assoc} = \frac{R}{[L]_o - \frac{R[P]_o}{1+R}} \tag{2.6b}$$

Using the approach outlined above, K_{assoc} can be readily determined from a nanoES-mass spectrum obtained from solutions containing known initial concentrations of protein and carbohydrate ligand. However, for the MS-derived values to be meaningful, the ratio of bound to unbound protein initially present in the nanoES solution must be conserved throughout the nanoES-MS experiment and reflected in the mass spectrum. As described earlier, the ratio may be influenced by solution and gas phase processes (*e.g.* increase in analyte concentration in the droplets resulting from solvent evaporation, changes in solution pH, formation of nonspecific complexes during the nanoES process and dissociation of the gaseous complex in the ion source). We have investigated the influence of these processes on binding affinity measurements in an attempt to identify the optimal experimental parameters/conditions for such measurements and the results of this study are described below.

2.3.2 scFv-αGal[αAbe]αMan Binding Affinity Measurements

Shown in Figure 2.3 is a typical nanoES mass spectrum obtained in positive ion mode for an aqueous solution containing scFv and αGal[αAbe]αMan with 1 mM ammonium acetate (pH 7). The dominant ions observed in the mass spectrum correspond to the protonated, unbound, scFv ion, $(scFv + nH)^{n+} \equiv scFv^{n+}$, and the protonated complex, $(scFv + Gal\alpha[Abe]Man + nH)^{n+} \equiv (scFv + \alpha Gal[\alpha Abe]\alpha Man)^{n+}$, with charge states of n = 9 – 11. From the expanded portion of the mass spectrum shown in Figure 2.3, it can be seen that there are two peaks corresponding to scFv ions. The higher mass peak coincides with the expected mass of the scFv, based on the DNA sequence of the plasmid used to express the protein in *E. coli*. A second peak, which is believed to correspond to a modified scFv, 18 Da lower in mass, is also present. At this time, the nature and site of the modification are not known. However, the modification does not influence the interaction between the protein and the carbohydrate ligand and is, therefore, likely remote from the binding site. The attachment of one or more Na^+ and K^+ ions to the protein and complex are also commonly observed, particularly at high solution flow rates, > 50 nL/min.

From the nanoES-mass spectra, such as the one shown in Figure 2.3, *R* values can be calculated from the abundance of bound and unbound protein at each of the observed charge states. One might expect that the *R*'s would be independent of charge-state; however, this was not found to be the case. Shown in Figure 2.4 are the charge-state dependent ratios (R_n) calculated from the intensities of the +9 to +11 ions (*i.e.* R_{+9}, R_{+10}, R_{+11}) from mass spectra acquired with two different nanoES tips for a solution containing

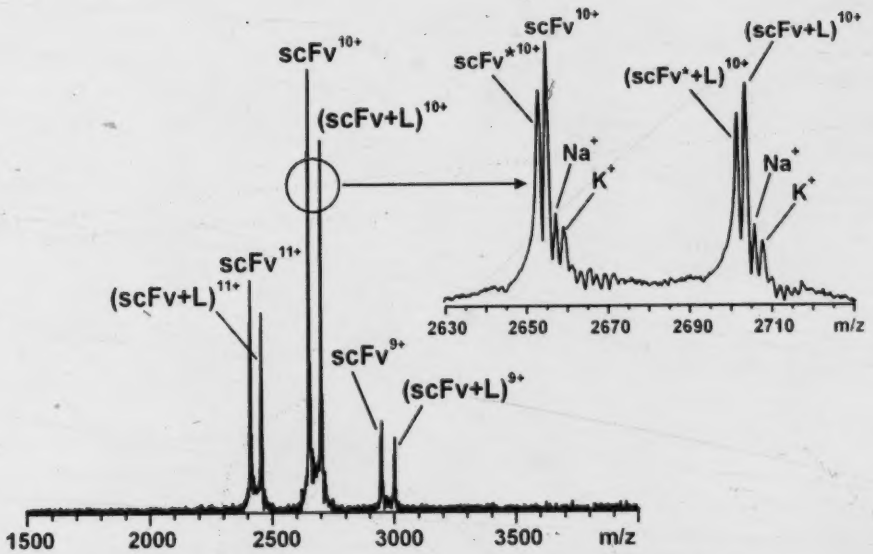


Figure 2.3 A typical nanoES mass spectrum obtained in positive ion mode for an aqueous solution containing equimolar (6.3 μM) scFv and its native trisaccharide ligand (L) αGal[αAbe]αMan with 1 mM ammonium acetate (pH 7). Insert shows expanded *m/z* region for the +10 charge state; the modified scFv is labeled with an asterisk (*).

equimolar scFv and αGal[αAbe]αMan (6.3 μM), based on eq. 2.3. It can be seen that the R_n values determined from a given mass spectrum differ by as much as 40 % in some cases. Also, for a given n, the R_n's values are not constant, but vary for individual measurements obtained with the same tip, and between tips. Furthermore there is no correlation between the magnitude of R_n and n. A survey of ES-MS data for a variety of protein-ligand complexes reveals that this is a general phenomenon [14, 15, 21]. However, to our knowledge, its origin has not been previously explained. Interestingly, when the R values are calculated from individual mass spectra as the sum of the intensities of complex ions, at all charge states, divided by the sum of the intensities of

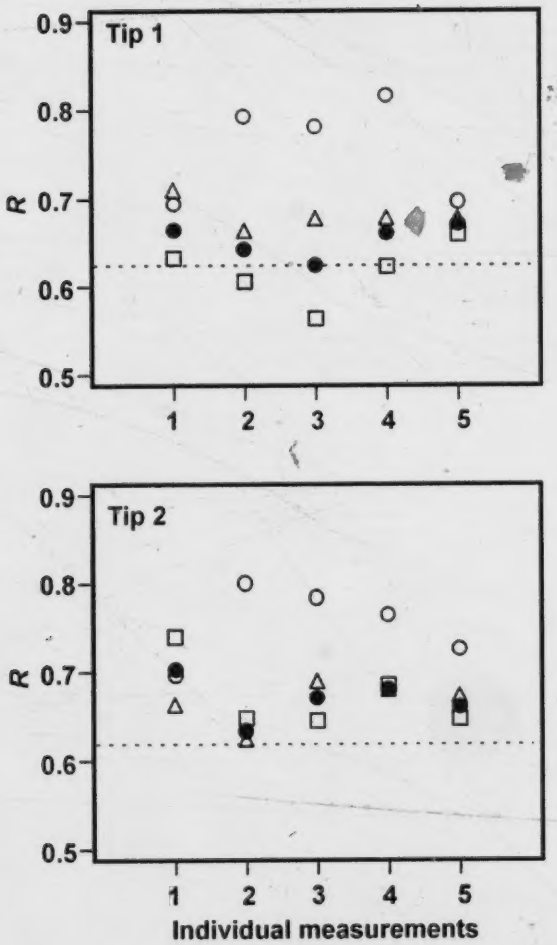


Figure 2.4 Distribution of R_n obtained from five individual measurements with two different nanoES tips (Tip 1 and Tip 2) for equimolar (6.3 μM) solutions of scFv and αGal[αAbe]αMan: ○, R_{+9}; □, R_{+10}; Δ, R_{+11}; ●, R calculated with eq 2.7. Dashed line indicates the calculated R value (≡ R_{ITC}) of 0.62.

the protein ions, at all charge states, the values are found to be in reasonable agreement with the expected ratio, R_{ITC}, which is calculated from the ITC-derived K_{assoc} at 298K and

the known initial solution concentrations, Figure 2.4. Since the signal intensity in FT-ICR/MS is proportional to the abundance and the charge state of the ion [29] the value of R was calculated using the charge-normalized ion intensities:

$$R = \frac{[PL]_{equil}}{[P]_{equil}} = \frac{\sum_{n}(I_{(P\bullet L)^{n+}}/n)}{\sum_{n}(I_{(P)^{n+}}/n)} \tag{2.7}$$

The agreement between the R values, calculated using eq. 2.7, and R_{ITC} suggests that the differences in the R_n's, for a given measurement, are due in large part to the differential partitioning of charge between the bound and unbound protein ions, which results in different charge state distributions for the two species. Although the mechanism by which gaseous protein ions are produced by ES and nanoES is still a matter of debate, there is experimental evidence that the charge residue model (CRM), wherein the initial droplets undergo successive fission events leading ultimately to multiply charged nanodroplets containing single analyte molecules, operates for macromolecules. These nanodroplets subsequently undergo solvent evaporation to yield multiply charged gaseous ions. According to this mechanism, the maximum charge state of the gaseous protein is determined by the charge supported by a nanodroplet of approximately the same size as the ion [30]. Statistical fluctuations in the number of charges carried by the droplets, which ultimately lead to gaseous protein and protein-ligand complexes, would result in differing charge state distributions for the two species and explain the differences in the R_n values. In addition, charge transfer reactions occurring in the ion source may also influence the final charge distribution of the protein and protein-ligand ions.

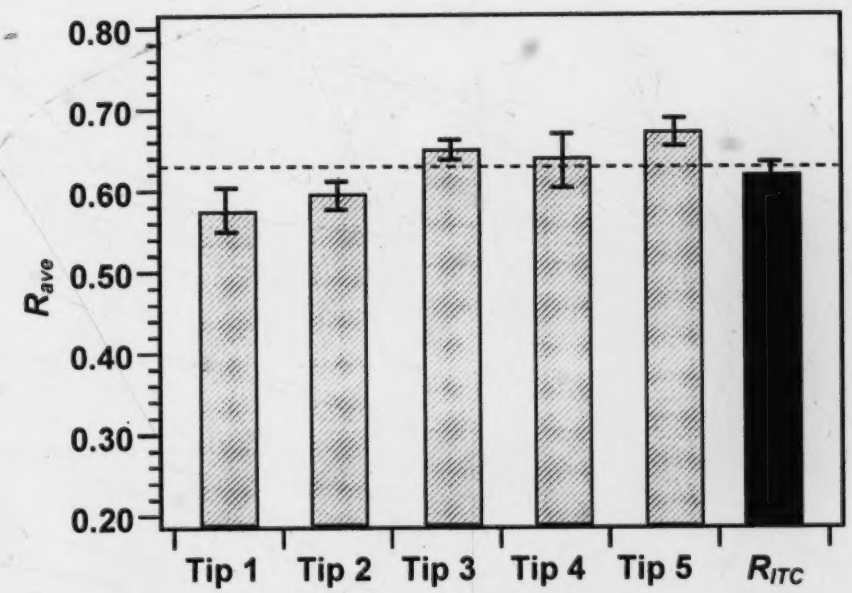


Figure 2.5 Comparison of R_{ave}, measured with five nanoES tips for equimolar (6.3 μM) solutions of scFv and αGal[αAbe]αMan, with the calculated R_{ITC} value. Error bars correspond to values of one standard deviation. Dashed line corresponds to the average R_{ave} value determined from the five sets of measurements (standard deviation is 0.27).

These fluctuations in the abundance ratio R can be minimized by making multiple measurements and averaging. The precision of the R_{ave} values, where R_{ave} is the average R value determined from multiple measurements (typically ≥ 5), obtained with different tips, was found to be quite good. Shown in Figure 2.5 are R_{ave} values determined with five nanoES tips. The R_{ave} values are similar, ranging from 0.57 to 0.67, with an average value of 0.63. This value and the corresponding K_{assoc} of $(1.63 \pm 0.27) \times 10^5$ M^{-1} are in excellent agreement with the R_{ITC} value of 0.62 and K_{assoc} of $(1.6 \pm 0.2) \times 10^5$ M^{-1}. As

described in a proceeding section, similar results were obtained for solutions containing different initial analyte concentrations.

The agreement between the MS- and ITC-derived K_{assoc} values for the scFv•αGal[αAbe]αMan complex indicates that the original equilibrium distribution of bound and unbound scFv in the nanoES solution is not altered by the nanoES/desolvation process. This may seem surprising given the fact that the concentration of analyte will increase during the life of the droplet, due to solvent evaporation. However, this observation can be rationalized by comparing the association-dissociation kinetics of the scFv•αGal[αAbe]αMan complex, measured by SPR [31], with the predicted lifetime of the nanoES droplets.

The diameters (d) of the droplets produced by nanoES can be estimated using the following equation, which was proposed by de la Mora [32]:

$$d \approx (V_f^2 \rho / \gamma)^{\frac{1}{3}} \tag{2.8}$$

where ρ is the density of the solution, V_f is the volume flow rate and γ is the surface tension of the solvent (0.072 N/m for water). This equation has been shown to accurately predict the diameter of ES droplets, within a factor of 2, in the ~0.1 μm to ~100 μm range [33]. In the present work V_f was in the range of 20 to 50 nL/min. These flow rates translate to initial droplet diameters of 115 to 210 nm, which are similar to the value of 200 nm predicted by Wilm and Mann for droplets produced by microelectrospray [9].

The radius (r) of the nanodroplets that ultimately produce gaseous protein ions (according to the CRM mechanism) can be estimated from the Rayleigh equation [34]:

$$Z_r = \frac{8\pi}{e}(\gamma \varepsilon_o r^3)^{\frac{1}{2}} \tag{2.9}$$

where Z_r is the maximum number of charges that can be supported by the droplet. The most abundant charge states observed for the $scFv^{n+}$ and $(scFv + \alpha Gal[\alpha Abe]\alpha Man)^{n+}$ ions, *i.e.* n = 9-12, correspond to a nanodroplet radius of ~ 2 nm.

As described by Beauchamp and coworkers [35], the time (t) required for a change in droplet diameter due to solvent evaporation can be estimated from the following expression:

$$d_p^2 = d_o^2 + st \tag{2.10a}$$

where d_o is the initial droplet diameter, d_p is the droplet diameter as a function of d_o, t and the solvent parameters which are included in s. For water droplets at a temperature of 278 K, s has a value of $-1250\ \mu m^2\ s^{-1}$ [35]. The above equation can then be rewritten as:

$$d_p^2 = d_o^2 - 1250t \tag{2.10b}$$

Assuming Rayleigh fission does not occur, the time required for a droplet with an initial diameter of 210 nm to shrink to 4 nm is approximately 35 μs. This value corresponds to an upper limit of the lifetime of the droplet, since one or more fission events are expected. Explicit consideration of droplet fission in the lifetime calculations, using the procedures outlined by Kebarle [36] and Enke [37], is expected to yield even shorter lifetimes. The lifetime of the scFv•αGal[αAbe]αMan $_{(aq)}$ complex at 298 K, which can be calculated from the dissociation rate constant (k_{off}) of 0.24 s^{-1} [31], is 4.2 s, a value at least 10^5 times greater than the expected lifetime of the droplet. The lifetime of the unbound scFv is estimated to be ~ 0.3 s, based on an association rate constant (k_{on}) of 3.8 x 10^4 $M^{-1}s^{-1}$ [31] and a free ligand concentration of 10^{-4} M. From this analysis it is clear that kinetics for the association/dissociation reactions are slow compared to the time required to produce gaseous ions. Consequently, the nanoES process will not

significantly alter the original distribution of bound and unbound protein - a necessary condition for the determination of K_{assoc} in solution by nanoES-MS.

2.3.3 Influence of Experimental Conditions

The MS-derived K_{assoc} for the scFv•αGal[αAbe]αMan complex reported above was determined using experimental conditions that were optimized for the binding affinity measurements. Described below are the results of a systematic investigation into the influence of spray duration, analyte concentration and ion source conditions on the binding measurements that was performed to identify these optimal conditions.

2.3.3.1 Solution pH

It is well established that electrochemical reactions, which occur at the electrode in the nanoES tip, can alter the composition of the nanoES solution [19]. When performing nanoES on aqueous solutions in the positive ion mode, the dominant electrochemical reaction at chemically inert electrodes such as platinum (used in the present work) is the oxidation of H_2O leading to the production of H_3O^+ (eq 2.11). Due to the small solution volumes used in nanoES, the electrochemical production of H^+ can lead to a significant decrease in pH [19, 20].

$$2H_2O \rightarrow O_2 + 4H^+ + 4e^- \qquad (2.11)$$

Since the binding affinity of protein-carbohydrate complexes (and protein-ligand complexes in general) is normally quite sensitive to the solution pH, the electrochemical oxidation of H_2O and concomitant drop in pH may introduce a significant error into the affinity measurements. Shown in Figure 2.6a are values of *R* measured with a single

nanoES tip during a 28 min period for a solution containing 14.7 μM scFv and 14.0 μM αGal[αAbe]αMan. The *R*'s measured in the first 10 min are found to be in good agreement with the R_{ITC} value of 1.33. After 10 min of spraying, *R* decreased with increasing spray duration, reaching a value of 0.66 at 28 min. The drop of ~ 0.6 in *R*, which corresponds to a 63 % decrease in K_{assoc}, is attributed to a pH-induced reduction in the stability of the scFv•αGal[αAbe]αMan complex which is known to be significant at pH < 5.0. In fact, the pH-dependent binding of Se155-4 and αGal[αAbe]αMan allows for the purification of the scFv by affinity chromatography under very mild conditions, wherein a change in buffer pH from 7.2 to 4.5 results in full elution of adsorbed scFv. It has been suggested that protonation of the residue His 35H, which is located at the base of the binding pocket and interacts with the abequose residue, causes the loss of ligand binding at pH < 5 [25]. While the pH of the nanoES solution was not monitored during the acquisition of the mass spectra, it was measured before and after the experiment and found to have decreased from 7.0 to 5.9. Van Berkel and coworkers have shown that the solution at the end of the nanoES tip, where the droplets are formed, experiences a more significant drop in pH compared to the bulk solution [19]. Therefore, the actual pH of the nanoES droplets produced after approximately 30 min of continuous spraying may be substantially below 5.9. Further evidence for pH-induced structural changes in the scFv can be found in the change in the relative abundance of the +9 to +12 protein and complex ions with spray duration, see Figure 2.6b. It can be seen that there is a continuous increase in the relative abundance of the +11 and +12 charge states with spray duration. This increase could be the result of the change in the protonation state of the

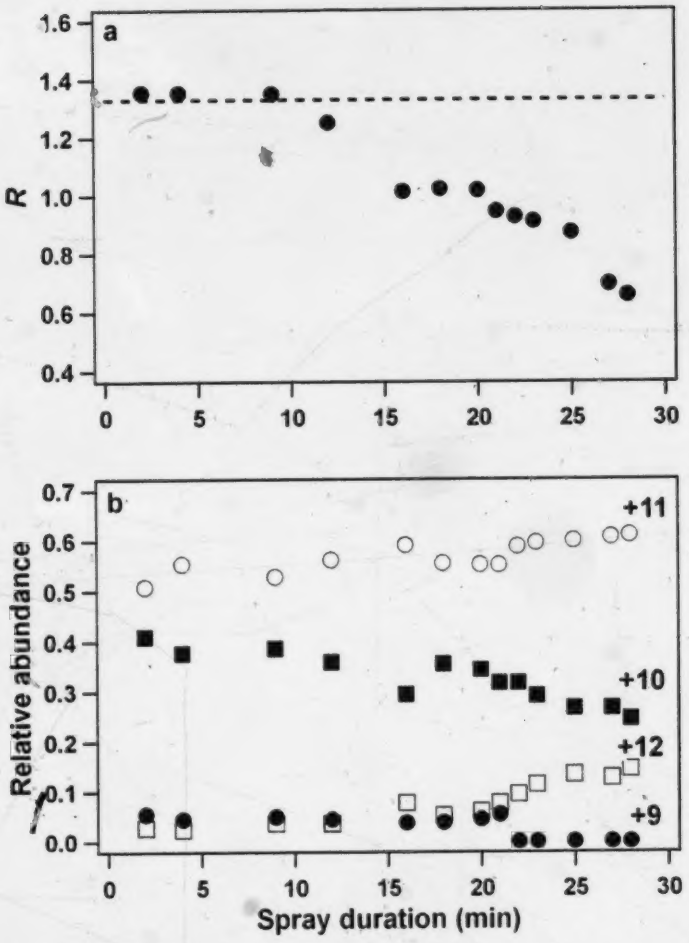


Figure 2.6 (a) Plot of R versus spray duration measured for a solution of scFv (14.7 μM) and αGal[αAbe]αMan (14.0 μM). Dashed line indicates the calculated R_{ITC} value of 1.33. (b) Plot of the relative abundance of the $scFv^{n+}$ and $(scFv + αGal[αAbe]αMan)^{n+}$ ions, at a given charge state, versus spray duration: ●, +9; ■, +10; ○, +11; □, +12.

protein in solution or some other structural change. Studies by Konermann and Douglas [38] have shown that the charge distribution observed for gaseous protein ions produced

by ES or nanoES is very sensitive to solution conformation, with the loss of native structure favouring higher charge states.

2.3.3.2 Analyte Concentration

MS-derived protein-ligand binding constants are normally determined, not from measurements of R at a single set of analyte concentrations, but rather from titration experiments where the concentration of one analyte (normally the protein) is fixed and the concentration of the other is varied [14, 15, 16b, 39, 40]. K_{assoc} can be extracted using nonlinear regression analysis of the experimentally determined concentration dependence of R, which is governed by the following expression:

$$R = \frac{1}{2}(-1 - K_{assoc}[P]_o + K_{assoc}[L]_o + \sqrt{(1 + K_{assoc}[P]_o - K_{assoc}[L]_o)^2 + 4K_{assoc}[L]_o}) \quad (2.12)$$

Ideally, R would be measured over a large range of analyte concentrations. However, in practice, there are limits to the analyte concentrations that can be investigated. Due to signal-to-noise ratio (S/N) considerations, measurements are normally limited to analyte concentrations that give R's of between ~ 0.1 and ~10. The formation of nonspecific complexes between analyte molecules during the nanoES process, which are commonly observed at high analyte concentrations (>10^{-4} M) [11, 21], obscure the solution binding stoichiometry and affinity and further restricts the range of analyte concentrations that may be investigated.

Shown in Figure 2.7 are results from a titration experiment performed with an initial scFv concentration fixed at 18.1 μM. At the lower ligand concentrations

investigated, there is excellent agreement between the R_{ave} and R_{ITC} values. However, at higher ligand concentrations (> 36 μM), the R_{ave} values are consistently greater than R_{ITC}. Inspection of the corresponding mass spectra revealed the presence of [scFv + 2(αGal[αAbe]αMan)]$^{n+}$ ions (see Figures 2.8a-c). Since the scFv has only one known carbohydrate binding site, the 1:2 complex must be an artifact of the nanoES process, resulting from nonspecific binding of a ligand to the 1:1 complex. If the formation of the nonspecific complexes is statistical (*i.e.* the bare scFv and 1:1 complex have equal probability of binding a nonspecific ligand), an artificially high R value should result, in agreement with the present experimental results. Although the nature of the interactions responsible for the nonspecific complexes is not known, binding likely results from

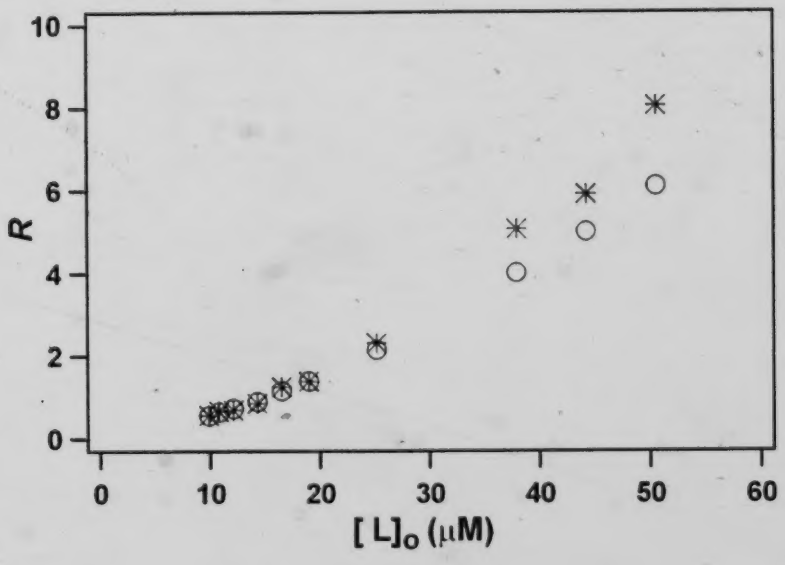


Figure 2.7 Plot of R versus the initial concentration of αGal[αAbe]αMan (≡ $[L]_o$) measured for a solution containing 18.1 μM scFv: ∗, MS-derived R_{ave}; ○, calculated R_{ITC}.

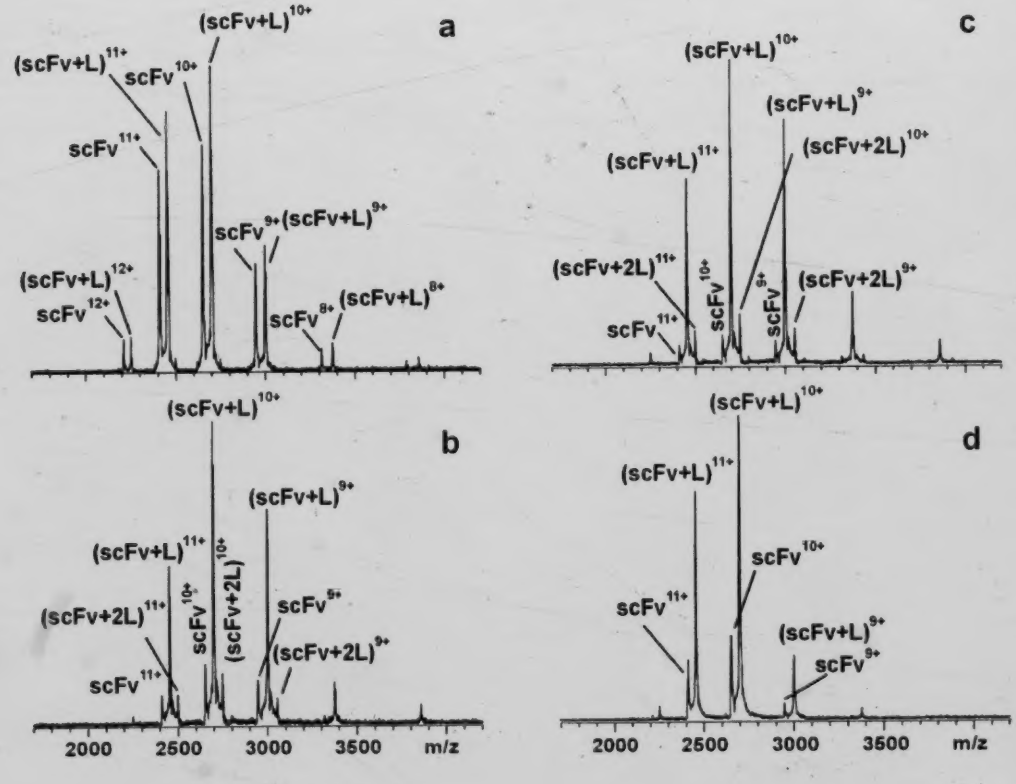


Figure 2.8 NanoES mass spectra obtained for aqueous solutions containing 18.1 μM scFv and (a) 19.0 μM; (b) 37.9 μM; (c) 50.6 μM αGal[αAbe]αMan (L); (d) equimolar (113 μM) scFv and αGal[αAbe]αMan (L).

random hydrogen bonds that form between the ligand and protein (or complex) during the final stages of desolvation. Interactions between the oligosaccharide ligand and protons on the surface of the protein may also help to stabilize the complex. Since the probability of forming nonspecific complexes is expected to increase with the number of free ligand molecule in the final nanodroplet, the occurrence of nonspecific binding

should increase with ligand concentration, which is consistent with the present experimental results, see Figures 2.8a-c.

Given that the nonspecific complexes were expected to be less stable than those originating from solution (*i.e.* specific complex), we considered the possibility of selectively removing the nonspecific ligand, by heating the ions in the ion cell, prior to acquiring the mass spectrum. To establish whether selective dissociation was feasible, the BIRD technique [41] was used to measure the dissociation kinetics and energetics of the nonspecific ligand in the $(scFv + \alpha Gal[\alpha Abe]\alpha Man)^{n+}$ complexes, where n = 10 – 11. Comparison of these results with data for complexes obtained from dilute solutions [42] revealed that the nonspecific ligand is kinetically more stable than the specific ligand, at least at these charge states, and that the dissociation activation energies are similar. These surprising results raise fundamental questions regarding the specificity of protein-ligand binding in the gas phase and indicate that the selective dissociation of nonspecific complexes is not possible in this case. Our laboratory extended the BIRD measurements to examine the influence of charge on the relative stability of specific and nonspecific complexes and these results will be described in Chapter 5.

Due to the tendency to form nonspecific complexes at [αGal[αAbe]αMan]/[scFv] concentration ratios ≥ 2, the traditional titration method was of limited use. Therefore, an alternative titration approach, employing equimolar solutions of protein and ligand, was used. With equimolar solutions, the concentration of free protein and free ligand at equilibrium are equivalent and $[L]_{equil}$ can be determined from R and $[P]_o$, according to eq. 2.13. A plot of $R(1+R)$, which can be determined directly from the mass spectrum, versus $[P]_o$ or $[L]_o$, should be linear with a slope equal to K_{assoc} (eq. 2.14).

$$[L]_{equil} = [P]_{equil} = [P]_o/(1+R) \tag{2.13}$$

$$K_{assoc} = \frac{[PL]_{equil}}{[P]_{equil}[L]_{equil}} = \frac{R}{[L]_{equil}} = \frac{R(1+R)}{[P]_o} \tag{2.14}$$

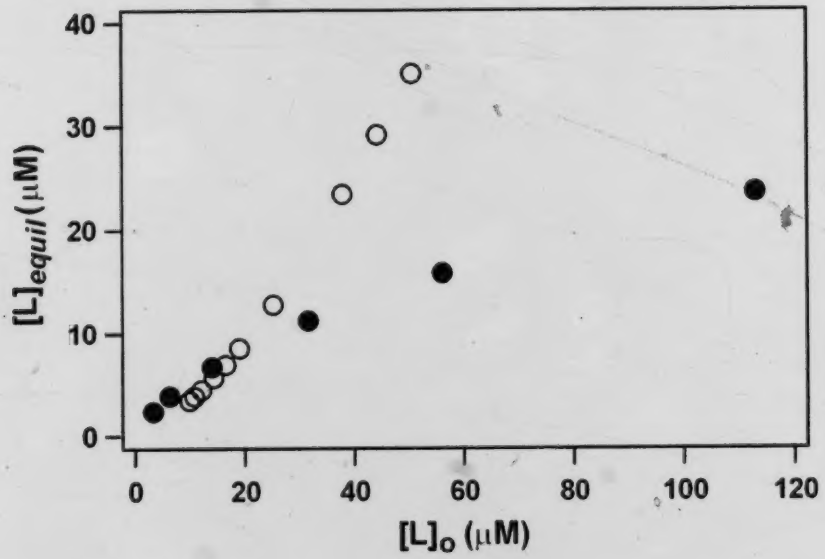


Figure 2.9 Plot of the dependence of the concentration of αGal[αAbe]αMan at equilibrium (≡ $[L]_{equil}$) on the initial concentration of αGal[αAbe]αMan (≡ $[L]_o$) for cases where $[scFv]_o$ fixed at 18.1 μM, ○ and $[scFv]_o = [L]_o$, ●.

The advantage of the equimolar approach, compared to the conventional titration method, can be seen in Figure 2.9, where the dependence of $[L]_{equil}$ on $[L]_o$, calculated for the cases where $[L]_o = [P]_o$ and $[P]_o$ = 18.1 μM, is shown. It can be seen that $[L]_{equil}$ increases much more slowly with $[L]_o$ when equimolar solutions are used. Since the probability of forming nonspecific complexes depends on the concentration of free ligand

in the nanodroplets, the use of equimolar solutions should minimize the occurrence of nonspecific complexes over a wide range of $[L]_o$.

Shown in Figure 2.10 is a plot of $R_{ave}(1+R_{ave})$ versus $[L]_o$, obtained for equimolar solutions of scFv and αGal[αAbe]αMan with initial concentrations ranging from 3.3 to 113 μM. From the slope of the linear titration plot a K_{assoc} of (1.70 ± 0.05) x 10^5 M^{-1} is obtained, which is in very good agreement with the ITC-derived value of (1.6 ± 0.2) x 10^5 M^{-1}. As can be seen in Figure 2.8d, nonspecific complexes were absent in the mass spectra, even at the highest concentrations investigated.

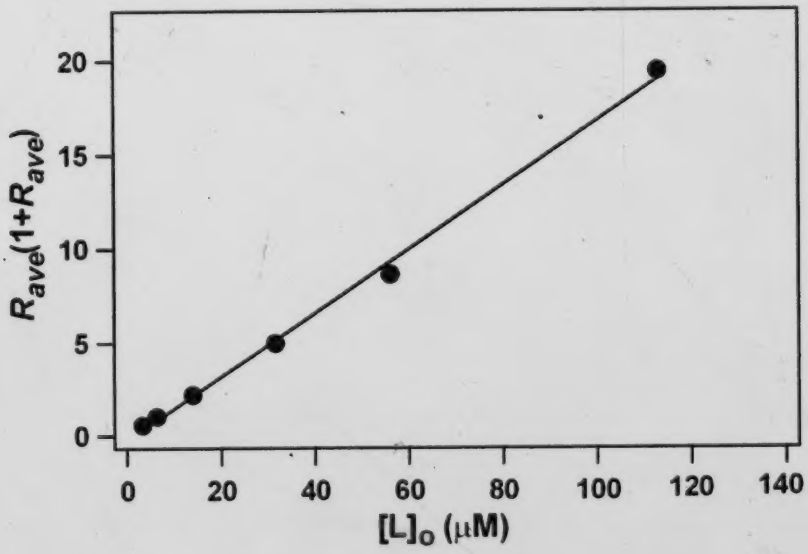


Figure 2.10 Plot of $R_{ave}(1+R_{ave})$ versus the initial concentration of αGal[αAbe]αMan (≡ $[L]_o$), measured for equimolar aqueous solutions of scFv and αGal[αAbe]αMan.

2.3.3.3 In-Source Dissociation

Collision-induced dissociation (CID) of the protein-ligand complex in the ion source of the mass spectrometer can introduce significant error into the binding affinity measurements. The extent to which collisional heating will influence the measurements depends on the configuration of the source and the size and gas phase stability of the complex. With the present apparatus, the internal energy of the gaseous ions sampled into the mass spectrometer was expected to be sensitive to the temperature of the inlet capillary, the nozzle-skimmer potential and the accumulation time in the hexapole. Varying the temperature of the sampling capillary between 305 and 340 K (external temperature) was found to have no noticeable effect on the relative abundance of the complex. The absence of dissociation in the sampling capillary can be explained by the fact that the nanoES ions being sampled into the capillary are likely extensively solvated. Consequently, the desolvation process will moderate the internal energy of the complex ions. A similar argument may be made to explain the fact that the *R* values were also found to be insensitive to modest changes (± 20 V) in the potential difference between the capillary and the skimmer (typically 52 and 5 V, respectively). In contrast, the accumulation time in the hexapole was found to have a significant effect on *R*. The pressure in the hexapole is not accurately known, but is believed to range from 10^{-5} to 10^{-3} torr. Acceleration of the ions by the rf field in the hexapole results in collisional-heating and, possibly the dissociation, of the gaseous $(scFv + \alpha Gal[\alpha Abe]\alpha Man)^{n+}$ ions. Shown in Figure 2.11a are values of *R* measured with accumulation times ranging from 1 to 7 s. With an accumulation time of 1 s, *R* is in good agreement with the expected value of 1.16. However, *R* decreases with longer accumulation times; by 60 % at 7 s (*R* =

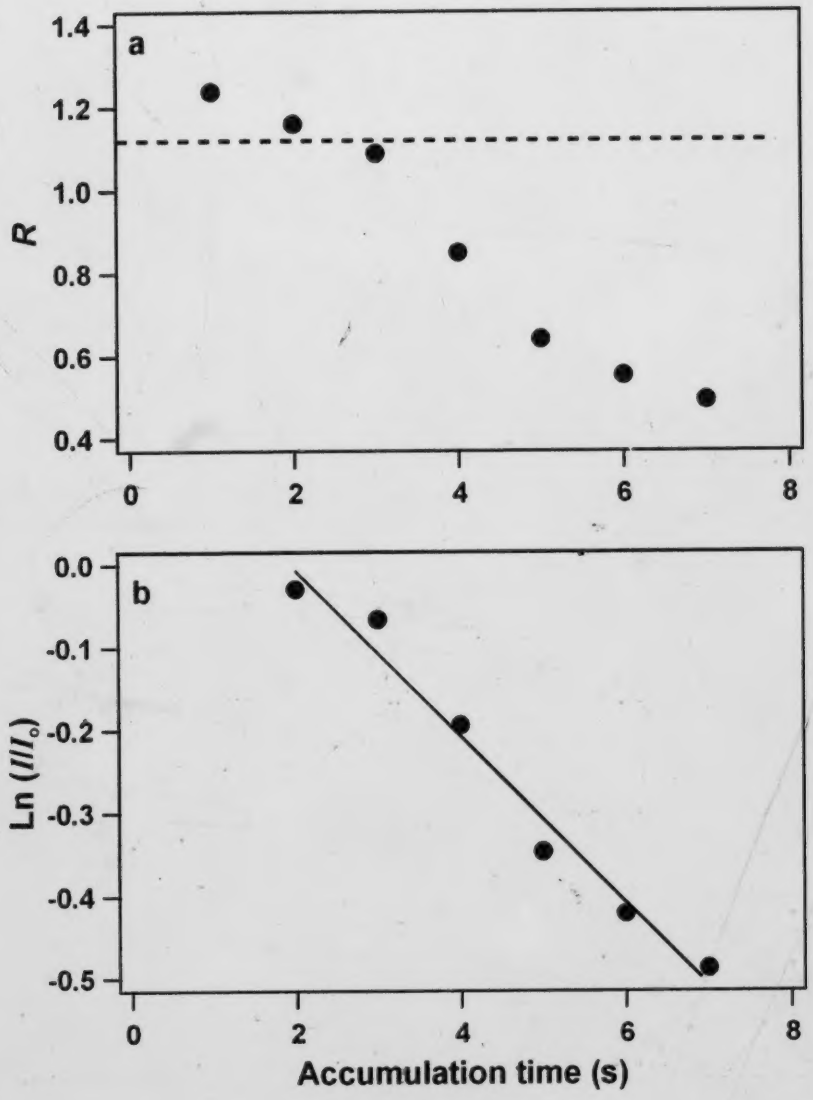


Figure 2.11 (a) Plot of R versus hexapole accumulation time, obtained from an aqueous solution of scFv (17 μM) and αGal[αAbe]αMan (L) (15 μM). Dashed line corresponds to the calculated R_{ITC} value; (b) Plot of $\mathrm{Ln}(I/I_o)$ versus accumulation time, where $I = \sum[I(\mathrm{scFv}^{n+})/n]$ and $I_o = \sum[(I(\mathrm{scFv}^{n+}) + I(\mathrm{scFv} + \alpha\mathrm{Gal}[\alpha\mathrm{Abe}]\alpha\mathrm{Man})^{n+})/n]$ measured with an accumulation time of 1s.

0.49). This drop in R translates to a decrease of 63 % in K_{assoc}. The decrease in R is attributed to CID of the (scFv + αGal[αAbe]αMan)$^{n+}$ ions while being stored in the hexapole, a phenomenon that has been previously described by Hofstadler and coworkers [43].

Given the continuous nature of the accumulation process, which leads to a distribution of ion residence times in the hexapole, the internal energy distribution of the ions is difficult to accurately characterize. Nevertheless, it is possible and insightful to estimate the *effective* temperature (T_{eff}) of the ions, which reflects their average internal energy, from the change in the relative abundance of complex ions with accumulation time. Shown in Figure 2.11b is a plot of $\ln(I/I_o)$ versus accumulation time, where I corresponds to the observed intensity of complex ions and I_o corresponds to the total ion intensity measured at an accumulation time of 1 s. The slope of the plot corresponds to an average dissociation rate constant of 0.11 s^{-1}. Comparison of this rate constant with the Arrhenius plots measured for the dissociation of the (scFv + αGal[αAbe]αMan)$^{n+}$ ions, where n = 9-11 [42, 44], indicates that ions stored in the hexapole for more than 2 s reach an *effective* temperature of ~ 415 K. These results clearly indicate that, in cases where the gaseous complexes are susceptible to dissociation at relatively low temperatures, short accumulation times, which minimize the extent of collisional-heating, are essential for obtaining reliable binding constants. However, spectra acquired with short accumulation times, < 1 s, generally suffer from poor S/N. An accumulation time of 1.5 s was used for the determination of the K_{assoc} values reported in this work.

2.3.4 Binding Affinities for scFv and Some Related Carbohydrate Ligands

To demonstrate that the experimental conditions identified above were generally suitable for protein-carbohydrate binding affinity determinations, K_{assoc} was measured for the scFv and three structurally-related ligands, αAbe(2-O-CH_3-αMan), βGlc [αAbe]αMan and αGlcαGlcαGal[αAbe]αMan. Association constants for IgG and these three ligands have been determined by ITC and found to range from 3 x 10^4 to 3 x 10^5 M^{-1} [27]. Furthermore, the stabilities of these complexes in the gas phase are quite different. Using the BIRD technique, the +10 complex of the pentasaccharide was found to be unreactive at temperatures up to 450 K, while the complex of αAbe(2-O-CH_3-αMan) is significantly more reactive (lifetime ≤ 1 s at 400K) and, therefore, more susceptible to dissociation in the source.

Using the equimolar titration approach, K_{assoc} for the complexes of the three ligands were determined and the values are listed in Table 2.1. Despite the range of gas and solution phase stabilities, the MS-derived binding constants are found to be in good agreement with the values determined by ITC. These results suggest that the experimental protocol described above is generally suitable for the determination of K_{assoc} for protein-carbohydrate complexes.

2.3.5 Competitive Binding Experiments

A unique feature of the MS-based approach for determining protein-ligand binding affinities is the ability to perform simultaneous measurements on multiple

analytes, provided that they have different molecular weights. This can be accomplished with a competition type experiment, performed with solutions containing a single protein and a mixture of ligands, where the total protein concentration is equal to or less than the total ligand concentration [14].

Table 2.1 Comparison of association constants (K_{assoc}) for scFv and its carbohydrate ligands determined by nanoES-FT-ICR/MS (MS) and isothermal titration calorimetry (ITC).

Ligand	K_{assoc} x 10^{-5} M^{-1} (MS)	K_{assoc} x 10^{-5} M^{-1} (ITC)[a]
αGal[αAbe]αMan	1.70 ± 0.05	1.6 ± 0.2
αAbe(2-O-CH_3-αMan)	1.50 ± 0.20	1.43 ± 0.05
βGlc[αAbe]αMan	0.15 ± 0.04	0.30 ± 0.14
αGlcαGlcαGal [αAbe]αMan	5.30 ± 0.50	3.81 ± 0.13

a. Bundle, D. R. unpublished data.

Shown in eq. 2.15a-n are the association equilibria involving a single protein (P) with a series of ligands (L_i, where i = 1, 2,..., n).

$$P + L_1 \leftrightarrow P{\bullet}L_1 \quad (2.15a)$$

$$P + L_2 \leftrightarrow P{\bullet}L_2 \quad (2.15b)$$

$$\vdots \quad \vdots \quad \vdots$$

$$P + L_n \leftrightarrow P{\bullet}L_n \quad (2.15n)$$

If K_{assoc} for one of the ligands is known, the corresponding protein-ligand complex can be used as an internal reference for evaluating the binding affinities for the other ligands. The approach is illustrated below for the determination of K_n, in the case where K_1 is

known. As shown in eq. 2.16, the ratio K_1/K_n depends on the ion intensity ratio of complexes (i.e. $I(PL_1)/I(PL_n)$) and the ratio of free ligand concentrations:

$$\frac{K_1}{K_n} = \frac{[PL_1][L_n]}{[PL_n][L_1]} = \frac{I(PL_1)[L_n]}{I(PL_n)[L_1]} \qquad (2.16)$$

As described previously by Roepstorff and coworkers [16a], the calculation is simplified when using equimolar solutions of protein and ligand. In this case, the ratio of K_1/K_n can be calculated simply from $I(PL_1)/I(PL_n)$ and the relative abundance of the PL_n complex (*i.e.* $R(PL_n)$, see eq. 2.17):

$$[L_n] = [P]_o[1 - I(PL_n)/(I(P) + \Sigma I(PL_n))] = [P]_o[1 - R(PL_n)] \qquad (2.17)$$

$$\frac{K_1}{K_n} = \frac{I(PL_1)[1 - R(PL_n)]}{I(PL_n)[1 - R(PL_1)]} \qquad (2.18)$$

Shown in Figure 2.12 are nanoES mass spectra collected from equimolar solutions containing the scFv and αAbe(2-O-CH_3-αMan), αGal[αAbe]αMan, and αGlcαGlcαGal[αAbe]αMan (6.1 and 14.8 μM). From the relative abundance of the three complexes, the order of the binding affinities can be immediately established: αGlcαGlcαGal[αAbe]αMan > αGal[αAbe]αMan > αAbe(2-O-CH_3-αMan), which is consistent with the order of affinities determined by ITC. Shown in Table 2.2 are the binding constants for αGlcαGlcαGal[αAbe]αMan and αAbe(2-O-CH_3-αMan), calculated from the measured ion abundance and a K_{assoc} of 1.70 x 10^5 M^{-1} for αGal[αAbe]αMan. It can be seen that there is very good agreement between the MS-derived constants determined from the competitive binding experiments and the ITC-derived values.

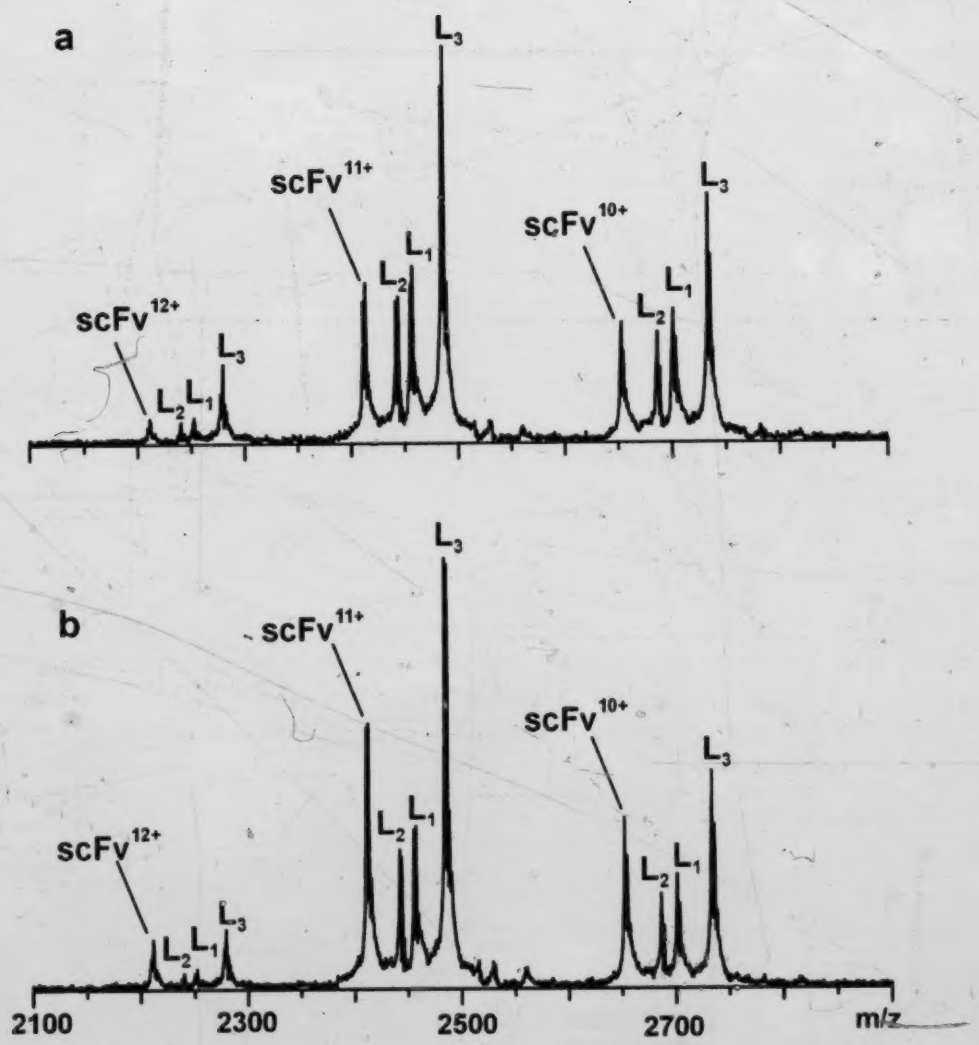


Figure 2.12 NanoES mass spectra obtained for equimolar aqueous solutions containing scFv, αGal[αAbe]αMan (L_1), αAbe(2-O-CH_3-αMan) (L_2), and αGlcαGlcαGal[αAbe]αMan (L_3): (a) 14.8 μM; (b) 6.1 μM.

Table 2.2 Association constants (K_{assoc}) for scFv and its carbohydrate ligands determined by nanoES-FT-ICR/MS (MS) and competition experiments performed at two different analyte concentrations (6.1 and 14.8 μM) and by isothermal titration calorimetry (ITC).

Ligand	K_{assoc} x 10^{-5} M^{-1} [a] (MS) 6.1 μM	K_{assoc} x 10^{-5} M^{-1} [a] (MS) 14.8 μM	K_{assoc} x 10^{-5} M^{-1} [b] (ITC)
αAbe(2-O-CH_3-αMan)	1.41 ± 0.08	1.38 ± 0.09	1.43 ± 0.05
αGlcαGlcαGal[αAbe] αMan	5.68 ± 0.28	5.10 ± 0.45	3.81 ± 0.13

a. The scFv•αGal[αAbe]αMan complex, with an MS-derived K_{assoc} of 1.70 x 10^5 M^{-1}, was used as the internal reference for the competition experiments.

b. Bundle, D. R. unpublished data.

2.4 Conclusions

The influence of analyte concentration, solution pH and in-source dissociation on the determination of K_{assoc} for the scFv•αGal[αAbe]αMan complex by nanoES-FT-ICR/MS has been investigated. From titration experiments, wherein the protein concentration was fixed and the ligand concentration was varied, it was found that the nanoES process leads to the formation of nonspecific complexes at modest free ligand concentrations, > 20 μM. The formation of nonspecific complexes leads to an artificially high K_{assoc}. The influence of nonspecific binding can be minimized, however, by employing an equimolar titration approach. K_{assoc} was found to decrease with spray duration due to the electrochemical production of the hydronium ion at the nanoES electrode. Limiting the spray duration to < 10 min minimized the influence of pH-induced loss of binding. Extensive collision-induced dissociation of the gaseous complex ions occurred during accumulation of ions in the trapping hexapole of the ion source.

From the change in the extent of dissociation with accumulation time, an *effective* ion temperature of ~415 K was determined. To minimize the extent of collision-induced dissociation, a short accumulation time, 1.5 s, was adopted. The MS-derived K_{assoc} measured for scFv•Galα[Abe]Man using these optimal conditions was found to be in excellent agreement with the previously reported value obtained using isothermal titration microcalorimetry. To demonstrate that the experimental conditions could be generally applied to protein-carbohydrate complexes, binding affinity measurements were performed on the scFv with a series of carbohydrate ligands. Although these complexes exhibit a range of solution and gas phase stabilities, the MS- and ITC-derived K_{assoc}'s are in excellent agreement. Future studies will focus on the application of the nanoES-MS approach to determine K_{assoc} for other classes of non-covalent protein complexes, including protein dimers and multimers.

2.5 Literature Cited

(1) Lis H.; Sharon N. *Chem. Rev.* **1998**, *98*, 637-674.

(2) Helnius, A.; Aebi, M. *Science* **2001**, *23*, 2364-2369.

(3) Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A. *Science* **2001**, *23*, 2370-2376.

(4) (a) Bundle, D. R. *Methods Enzymol.* **1994**, *247*, (Chap. 20) 288-305; (b) Bundle, D. R. In *Carbohydrates*, Hecht, S., Ed., Oxford University Press: Oxford, **1998**. pp. 370-440.

(5) Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555-578.

(6) (a) Zhang B. Y.; Chan N.; Schriemer D. C.; Palcic M. M.; Hindsgaul O. *Glycobiology* **1999**, *9*, 14; (b) Schriemer D. C.; Hindsgaul O. *Comb. Chem. High T. Scr.* **1998**, *1*, 155-170.

(7) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1-27.

(8) Wilm M.; Mann M. *Anal. Chem.* **1996**, *68*, 1-8.

(9) Wilm M. S.; Mann M. *Int. J. Mass Spectrom. Ion processes* **1994**, *136*, 167-180.

(10) Ganem, B.; Li, Y.-T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 7818-7819.

(11) Van Dongent, W. D.; Heck, A. J. R. *Analyst* **2000**, *125*, 583-589.

(12) Siebert H.-C.; Lü S.-Y.; Frank M.; Kramer J.; Wechselberger R.; Joosten J.; André S.; Rittenhouse-Olson K.; Roy R.; von der Lieth C.-W.; Kaptein R.; Vliegenthart J. F. G.; Heck A. J. R.; Gabius H.-J. *Biochemistry* **2002**, *41*, 9707-9717.

(13) Ayed, A.; Krutchinsky, A. N.; Ens, W.; Standing, K. G.; Duckworth, H. W. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 339-344.

(14) Loo, J. A.; Hu, P.; McConnell, P.; Mueller, W. T.; Sawyer, T. K.; Thanabal, V. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 234-243.

(15) Greig, M. J.; Gaus, H.; Cummins, L. L.; Sasmor, H.; Griffey, R. H. *J. Am. Chem. Soc.* **1995**, *117*, 10765-10766.

(16) (a) Jorgensen, T. J. D.; Roepstorff, P.; Heck, A. J. R. *Anal. Chem.* **1998**, *70*, 4427-4432; (b) Sannes-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264-271.

(17) Griffey, R. H.; Sannes-Lowery, K. A.; Drader, J. J.; Mohan, V.; Swayze, E. E.; Hofstadler, S. A. *J. Am. Chem. Soc.* **2000**, *122*, 9933-9938.

(18) Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605-611.

(19) Van Berkel, G. J.; Asano, K. G.; Schnier, P. D. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 853-862.

(20) Van Berkel, G. J.; Zhou, F.; Aronson, J. T. *Int. J. Mass Spectrom. Ion processes* **1997**, *162*, 55-67.

(21) (a) Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. *J. Am. Chem. Soc.* **1996**, *118*, 8646-8653; (b) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(22) Collings, B. A.; Douglas, D. J. *J. Am. Chem. Soc.* **1996**, *118*, 4488-4489.

(23) (a) Toone, E. J. *Curr. Opin. Struct. Biol.* **1994**, *4*, 719-728; (b) Dam, T. K.; Roy, R.; Das, S. K.; Oscarson, S.; Brewer, C. F. *J. Biol. Chem.* **2000**, *275*, 14223-14230.

(24) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S.-J.; MacKenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6423-6427.

(25) Deng, S.-J.; MacKenzie, C. R.; Sadowska, J.; Mickniewicz, J.; Young, M.; Bundle, D. R.; Narang, S. A. *J. Biol. Chem.* **1994**, *269*, 9533-9538.

(26) Bundle D. R. unpublished data.

(27) Bundle D. R.; Eichler, E.; Gidney, M. A. J.; Meldal, M.; Ragauskas, A.; Sigurskjold, B. W.; Sinnott, B.; Watson D. C.; Yaguchi, M.; Young, N. M. *Biochemisty* **1994**, *33*, 5172-5182.

(28) Felitsyn, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2001**, *73*, 4647-4661.

(29) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.

(30) de la Mora, J. F. *Anal. Chim. Acta.* **2002**, *406*, 93-104.

(31) MacKenzie, C. R.; Hirama, T; Deng S.-J.; Bundle, D. R.; Narang, S. A.; Young, N. M. *J. Biol. Chem.* **1996**, *271*, 1527-1533.

(32) de la Mora, J. F.; Navascues, J.; Fernandez, F.; Rosell-llompar J. *J. Aerosol Sci.* **1990**, *21*, 5673-5676.

(33) de la Mora, J. F. *J. Fluid. Mech.* **1992**, *243*, 561-574.

(34) Rayleigh, L. *Phil. Mag.* **1882**, *14*, 184-186.

(35) Grimm R. L.; Beauchamp J. L. *Anal. Chem.* **2002**, *74*, 6291-6297.

(36) Kebarle, P; Tang L. *Anal. Chem.* **1993**, *65*, 972A-986A.

(37) Cech, N. B.; Enke, C. G. *Anal. Chem.* **2001**, *73*, 4632-4639.

(38) Konermann, L; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1248-1254.

(39) Hirose, K. *Journal of Inclusion Phenomena and Macrocyclic Chemistry* **2001**, *39*, 193-209.

(40) (a) Lim, H.–K.; Hsieh, Y. L.; Ganem, B.; Henion, J. *J. Mass Spectrom.* **1995**, *30*, 708-714; (b) Griffey, R. H.; Hofstadler, S. A., Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 10129-10133.

(41) (a) Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644; (b) Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194-197.

(42) Kitova, E. N.; Bundle, D. R.; Klassen J. S. *J. Am. Chem. Soc.* **2002**, *124*, 13980-13981.

(43) Hofstadler, S. A.; Sannes-Lowery, K. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1-9.

(44) Klassen, J. S. unpublished results.

Chapter 3

Bioactive Recognition Sites May Not Be Energetically Preferred in Protein-Carbohydrate Complexes in the Gas Phase*

3.1 Introduction

Noncovalent solute-solute interactions and solvent effects govern affinity and specificity in biological recognition. Studies of desolvated biomolecular complexes may afford insight, hitherto unavailable, into their intrinsic binding affinity. To date, the limited structural studies of gaseous biomolecular complexes have focused primarily on the extent to which the structure of specific complexes, pre-formed in solution and transferred to the gas phase by ES or nanoES, is preserved [1, 2]. A related and, as yet unexplored, question is how the method of preparation influences the structure and stability of gaseous complexes.

Here, we describe results from thermal dissociation experiments performed on gaseous protein-carbohydrate complexes originating from nonspecific interactions during the nanoES process. The stability of the nonspecific complexes compared to the corresponding specific complex originating from interactions in solution provides the first evidence that a bioactive recognition site is not energetically preferred in the gas phase.

A version of this chapter has been published: Wang, W.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Chem.* **2003**, *125*, 13630-13631.

3.2 Experimental

3.2.1 Protein and Carbohydrate Ligand

The single chain fragment, scFv (MW 26 539 Da), of the monoclonal antibody Se155-4, was produced using recombinant technology [3]. The scFv was concentrated and dialyzed against deionized water using MICROSEP microconcentrators with a molecular weight cut-off of 10 kDa, and lyophilized prior to MS analysis. Bovine carbonic anhydrase II, CA (MW 29 089 Da) was purchased from Sigma (Canada). This compound was used without further purification. The trisaccharide ligand, αGal[αAbe]αMan, used in this work was provided by D. R. Bundle (Univ. of Alberta). The structure of αGal[αAbe]αMan is shown in Figure 2.1 of Chapter 2.

3.2.2 Mass Spectrometry

Gaseous protein-carbohydrate complexes were produced by nanoES performed on aqueous solutions containing 19 μM scFv and 19 or ≥ 38 μM trisaccharide ligand, αGal[αAbe]αMan (**1**), or 8 μM CA and ≥ 40 μM **1**, and 1 mM $CH_3CO_2NH_4$. NanoES tips were constructed from aluminosilicate capillaries (1 mm o.d., 0.68 mm i.d.), pulled to 4 - 7 μm o.d. at one end using a P-2000 puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution was established by applying a voltage of 800-1000 V to a platinum wire inserted inside the glass tip. The solution flow rate was typically 20 - 50 nL/min. The droplets and gaseous ions produced by nanoES were introduced into the vacuum chamber of an ApexII 47e Fourier-transform ion cyclotron resonance mass spectrometer (Bruker, Billerica MA) through a stainless steel capillary (i.d. 0.43 mm)

maintained at an external temperature of 66 °C. The ion/gas jet sampled by the capillary (52 V) was transmitted through a skimmer (4 V) and stored, in a hexapole. Ions were accumulated in the hexapole for 2 - 3 s, then ejected and accelerated (~ -2700 V) into a 4.7 tesla superconducting magnet, decelerated and introduced into the ion cell. The trapping plates of the cell were maintained at a constant potential of 1.3 V throughout the experiment. The typical base pressure for the instrument was ~ 5×10^{-10} mbar.

The temperature of the ion cell was controlled with two external flexible heating blankets placed around the vacuum tube in the vicinity of the cell. In a separate experiment, the temperature inside the cell, measured by a thermocouple placed temporarily inside the cell, was calibrated against the temperature measured by eight thermocouples placed on the outside of the vacuum tube. Using this approach, calibration plots were generated for cell temperatures ranging from 25 to 175 °C.

Data acquisition was controlled by an SGI R5000 computer running the Bruker Daltonics XMASS software, version 5.0. Mass spectra were obtained using standard experimental sequences with chirp broadband excitation. Isolation of the reactant ions for the BIRD experiments was achieved using a combination of single rf frequency and broadband rf sweep excitation. The isolated ions were stored inside the heated cell for varying reaction times prior to excitation and detection. The excitation pulse length was 15 μsec and the power of excitation pulse was varied to maximize the intensity of the ion signal. The time-domain spectra, consisting of the sum of 15 - 40 transients containing 128 K data points per transient, were subjected to one zero-fill prior to Fourier-transformation.

3.3 Results and Discussion

The carbohydrate-binding single chain fragment (scFv) of the monoclonal antibody Se155-4 and its native trisaccharide ligand, αGal[αAbe]αMan (**1**) [3], served as a model system for comparing the energetic and kinetic stability of a gaseous protein-ligand complex produced from specific interactions in solution and nonspecific interactions during the nanoES process. NanoES performed on an equimolar aqueous solutions of scFv and **1** yields predominantly the protonated $(scFv)^{n+}$ and $(scFv + \mathbf{1})^{n+}$ ions, where n = 9 – 11, see Figure 3.1a.

Using thermal dissociation experiments and functional group replacement, it has been shown [2] that one of the specific intermolecular hydrogen bonds, His^{101H} – OH_{ManC4}, is preserved in the gaseous $(scFv + \mathbf{1})^{n+}$ ions at n = 10. At n = 11, this specific hydrogen bond is absent. However, both the His^{101H} and Man C-4 OH groups contribute to the stability of the complex [4]. These results indicate that the specific interactions formed in solution impose constraints on the structure of the desolvated $(scFv + \mathbf{1})^{n+}$ ions. At concentration ratios [**1**]/[scFv] ≥ 2, $(scFv + 2(\mathbf{1}))^{n+}$ ions are also observed in the nanoES mass spectrum [5], see Figure 3.1b. Since the scFv has only a single binding site for **1**, the $(scFv + 2(\mathbf{1}))^{n+}$ ions must originate from nonspecific interactions between **1** and the specific scFv•**1** complex during the nanoES process. To distinguish these ligands, we will refer to them as specific ($\mathbf{1}_{sp}$) or nonspecific ($\mathbf{1}_{ns}$), *i.e.* $(scFv + 2(\mathbf{1}))^{n+} \equiv (scFv + \mathbf{1}_{sp} + \mathbf{1}_{ns})^{n+}$.

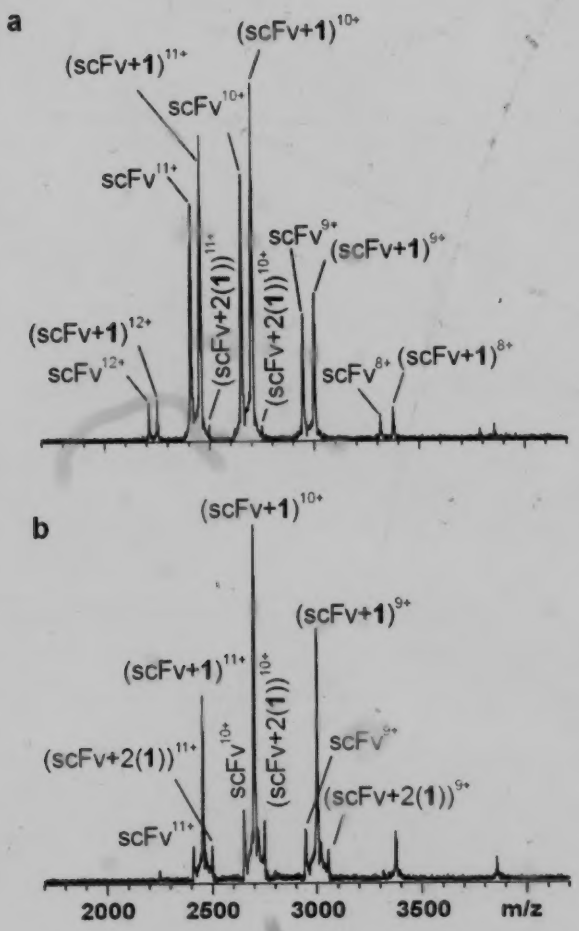


Figure 3.1 NanoES mass spectra obtained for buffered aqueous solutions (pH 7.0) containing scFv (19 μM) and its native trisaccharide ligand, αGal[αAbe]αMan (**1**), at (a) 19 μM and (b) 38 μM.

3.3.1 BIRD Pathways and Kinetics of scFv Nonspecific Complexes

Time-resolved thermal dissociation experiments were performed on the gaseous, protonated $(scFv + \mathbf{1})^{n+}$ and $(scFv + 2(\mathbf{1}))^{n+}$ ions, where n = 10, 11, produced by nanoES from solutions with [**1**]/[scFv] ≥ 2, to evaluate the kinetic and energetic stability of the

nonspecific complexes. The kinetic measurements were carried out using the BIRD technique [6] implemented with a modified FT-ICR/MS [7].

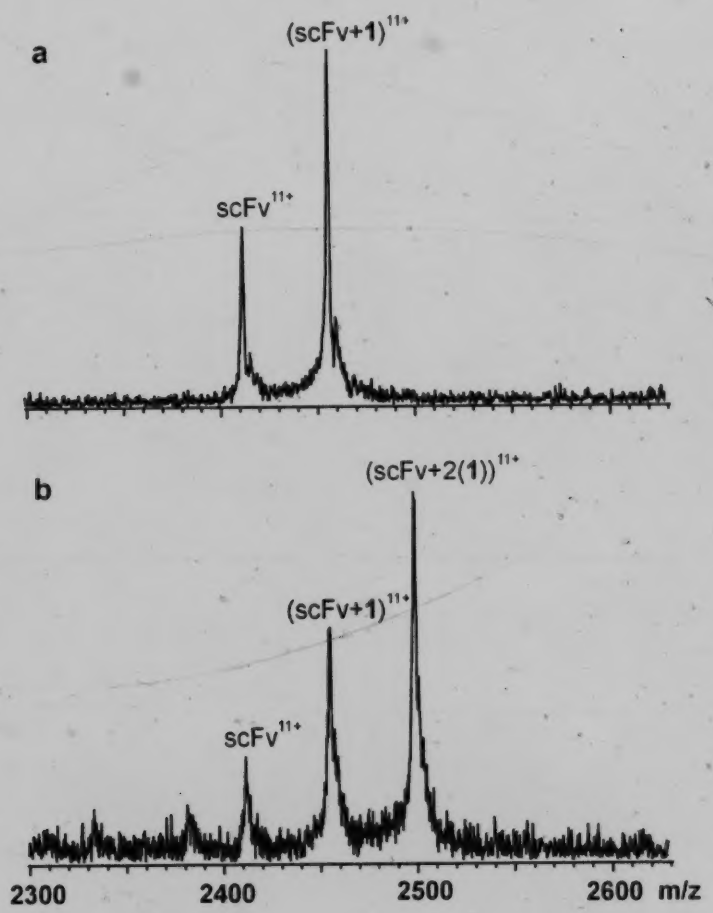


Figure 3.2 BIRD mass spectra obtained for the (a) $(scFv + \mathbf{1})^{11+}$, reaction time 1.8 s, reaction temperature149 °C, and (b) $(scFv + 2(\mathbf{1}))^{11+}$, 1.4 s, 149 °C.

At temperatures of 120 to 165 °C, dissociation of the complexes proceeds exclusively by the loss of neutral **1** (eq. 3.1). In the case of the $(scFv + 2(\mathbf{1}))^{n+}$ ions, the sequential loss of **1** is observed (eq. 3.2), as shown in Figure 3.2.

$$(scFv + 1)^{n+} \rightarrow scFv^{n+} + 1 \quad (3.1)$$

$$(scFv + 2(1))^{n+} \rightarrow (scFv + 1)^{n+} + 1 \rightarrow scFv^{n+} + 2(1) \quad (3.2)$$

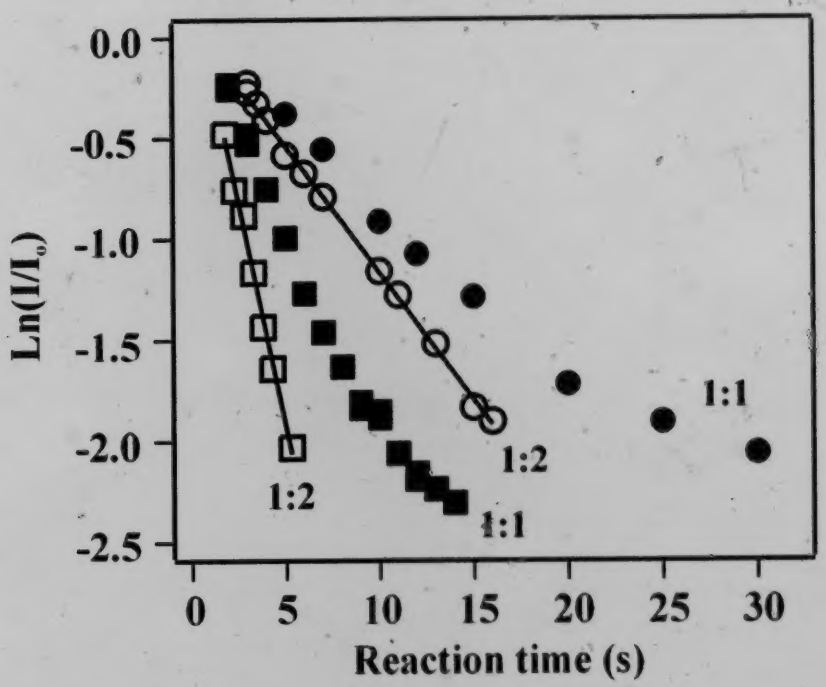


Figure 3.3 Plots of the natural logarithm of the normalized abundance (I/I_o) of the gaseous protonated complex $(scFv + 1)^{n+} \equiv 1{:}1$ (●, +10; ■, +11) and $(scFv + 2(1))^{n+} \equiv 1{:}2$ (○, +10; □, +11) ions versus reaction time measured at 142 °C. The rate constant (k_{obs}) for the loss of **1** from the protonated 1:2 complex was determined from a linear least-square fit.

Given that dissociation involves two parallel pathways (*i.e.* loss of $\mathbf{1}_{sp}$ or $\mathbf{1}_{ns}$), the measured rate constant (k_{obs}) is equal to the sum of rate constants for the loss of the $\mathbf{1}_{sp}$ (k_{sp}) and $\mathbf{1}_{ns}$ (k_{ns}), as described in eq. 3.3.

$$k_{obs} = k_{sp} + k_{ns} \quad (3.3)$$

The value of k_{ns} was determined from the measured k_{obs} and k_{sp}, which can be calculated at a given temperature from Arrhenius parameters for the dissociation of the (scFv + $\mathbf{1}_{sp}$)$^{n+}$ ions [2b,c]. In contrast, the natural log plots obtained for the dissociation of the (scFv + **1**)$^{n+}$ ions, produced at [**1**]/[scFv] $\geq$ 2, are nonlinear (Figure 3.3). The time dependence of the normalized abundance of the (scFv + **1**)$^{n+}$ ions is reasonably described by a double exponential function (see eq. 3.4, where f_{sp} and f_{ns} refer to the fraction of specific complexes and nonspecific complexes in the 1:1 complexes, respectively), which incorporates k_{sp} (known) and k_{ns}.

$$I/I_o = f_{sp}\exp(-k_{sp}t) + f_{ns}\exp(-k_{ns}t) \qquad (3.4)$$

This result indicates that two distinct species are present, *i.e.* (scFv + $\mathbf{1}_{sp}$)$^{n+}$ and (scFv + $\mathbf{1}_{ns}$)$^{n+}$. Importantly, the values of k_{ns} determined at a given temperature with these two approaches are identical, within experimental error. This suggests that the nonspecific interactions are insensitive to the presence of $\mathbf{1}_{sp}$.

3.3.2 Arrhenius Parameters

Arrhenius plots for the dissociation of the nonspecific protein-trisaccharide interactions were constructed from the temperature dependent rate constants and are shown in Figure 3.4. Also included are the Arrhenius plots for the dissociation of the corresponding (scFv + $\mathbf{1}_{sp}$)$^{n+}$ ions [2b,c]. The Arrhenius activation energies and pre-exponential factors, E_a and A, are listed in Table 3.1. The kinetic data obtained for the 1:1 and 1:2 complexes are consistent with the nonspecific ligand occupying a single site or, perhaps, several equivalent sites on the scFv. This result is surprising given that, by their very nature, nonspecific interactions are expected to lead to diverse structures. It may be

that the nonspecific complexes relax to a common, lowest energy structure (or several equivalent structures) after or in concert with the desolvation process. Alternatively, only a fraction of the nonspecific complexes produced by nanoES, those that are most stable, may contribute to the kinetic measurements. Complexes with less favourable modes of binding will have shorter lifetimes and may dissociate in the ion source.

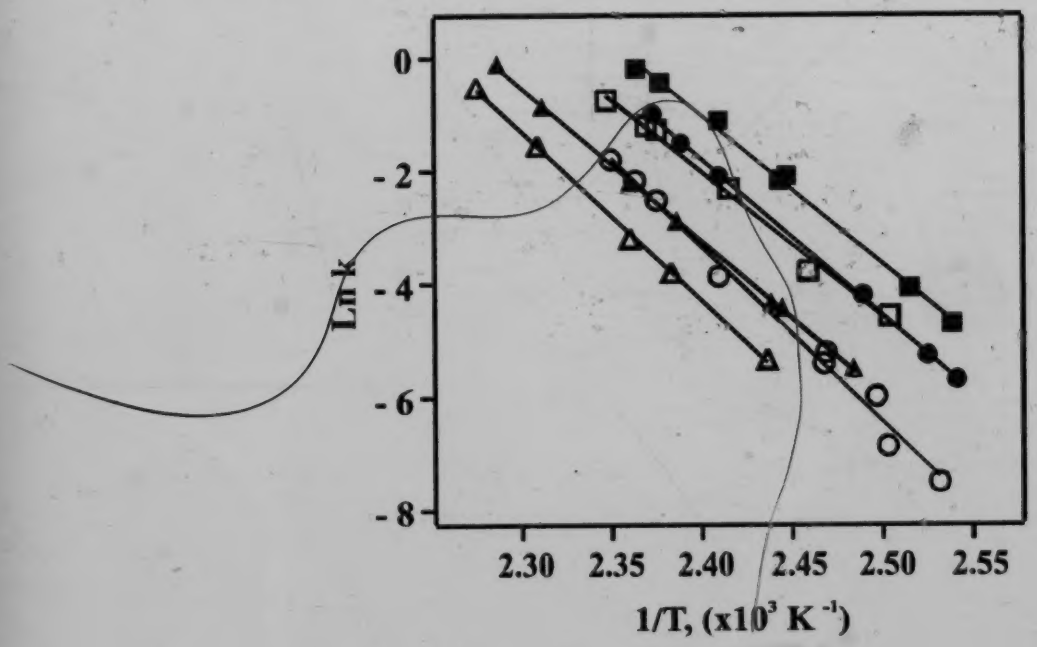


Figure 3.4 Arrhenius plots for the dissociation of $(\text{scFv} + \mathbf{1}_{ns})^{n+}$ (o, +10; □, +11), $(\text{scFv} + \mathbf{1}_{sp})^{n+}$ (●, +10; ■, +11) and $(\text{CA} + \mathbf{1})^{n+}$ (Δ, +10; ▲, +11) ions. Arrhenius parameters (E_a and A) for each complex were derived from a least-squares linear fit of each Arrhenius plot.

A comparison of the Arrhenius plots obtained for the dissociation of the specific and nonspecific complexes reveals that, over the temperature range investigated and at both charge states, the ions arising from nonspecific interactions are kinetically more stable (Figure 3.4). At the +11 charge state, the E_a's for the specific and nonspecific

complexes are similar, 51-52 kcal/mol and the greater kinetic stability of the nonspecific complex is entropic in origin. For the +10 ions, there is a striking difference in the E_a's, with the nonspecific complex being 7 kcal/mol more stable than the specific complex (61 versus 54 kcal/mol). It is also notable that the E_a for the +10 nonspecific complex is ~10 kcal/mol larger than for the +11 species.

Table 3.1 Arrhenius parameters determined for the dissociation of gaseous, protonated protein-trisaccharide complexes: $(P + L)^{n+} \rightarrow P^{n+} + L$.

P	L	Charge state	E_a [a] (kcal/mol)	A [a] (s^{-1})	$\Delta S^{\ddagger d}$ (cal/mol·K)
scFv	$\mathbf{1}_{sp}$	+10	54.3 ± 1.0[b]	$10^{27.7 \pm 0.6}$ [b]	65
	$\mathbf{1}_{ns}$		61.4 ± 2.4	$10^{30.8 \pm 1.3}$	80
	$\mathbf{1}_{sp}$	+11	52.1 ± 1.4[c]	$10^{26.9 \pm 0.8}$ [c]	61
	$\mathbf{1}_{ns}$		51.0 ± 2.3	$10^{25.9 \pm 1.2}$	57
CA	**1**	+10	59.7 ± 1.1	$10^{29.5 \pm 0.5}$	74
	1	+11	54.0 ± 0.5	$10^{27.0 \pm 0.3}$	62

a. The reported errors are values of one standard deviation. b,c. Arrhenius parameters taken from reference 2b and c, respectively. d. Values calculated at 415 K from the corresponding A-factors.

The enhanced energetic stability of the nonspecific +10 complex, compared to the +10 specific complex, is surprising. A possible explanation for this observation is that, in addition to a number of neutral hydrogen bonds, the $(\text{scFv} + \mathbf{1}_{ns})^{10+}$ ion is stabilized by an ionic hydrogen bond involving one of the protonated scFv residues. Based on thermochemical data measured for the interaction between H_2O and protonated peptides, which is the best available model, solvation of a protonated scFv residue by one of the

ligand OH groups can contribute as much as 15 kcal/mol to the stability of the complex [8]. An alternative explanation lies in conformational constraints imposed by the specific interactions in solution, which may "trap" the complex, preventing it from adopting an optimal structure for hydrogen bonding in the gas phase. Such constraints are expected to be absent in the case of $\mathbf{1}_{ns}$, allowing for more favourable interactions in the gaseous complex. Regardless of the nature of the interactions that stabilize the nonspecific +10 complex, the present results provide the first evidence that protein-ligand interactions remote from the bioactive recognition site can be energetically more favourable in the gas phase.

3.3.3 Arrhenius Parameters of CA Nonspecific Complexes

The propensity of **1** to form nonspecific interactions with proteins in the gas phase is further highlighted by the formation of the complex between bovine carbonic anhydrase II (CA) and **1**. CA, a Zn (II) metalloenzyme, does not interact specifically in solution with **1**, while the gaseous $(CA + \mathbf{1})^{n+}$ ions were readily observed resulting from nonspecific interactions during nanoES (Figure 3.5a). BIRD was performed on the $(CA + \mathbf{1})^{n+}$ complexes at charge states of 10 and 11. A typical BIRD spectrum, obtained for the +10 charge state at 165 °C, is shown in Figure 3.5b. It can be seen that the dissociation of the $(CA + \mathbf{1})^{n+}$ follows the same pathway as the $(scFv + \mathbf{1})^{n+}$, a neutral sugar lose. Kinetics plots obtained for the $(CA + \mathbf{1})^{10+}$ and $(CA + \mathbf{1})^{11+}$ are shown in Figure 3.6. The plots were found to exhibit good linearity (the correlation coefficient $\geq$ 0.99) over the reaction extent investigated. The temperature dependent rate constants for the loss of **1** from the $(CA + \mathbf{1})^{n+}$ were determined by using the simple first order reaction

with a single rate constant. Interestingly, the E_a's of 60 (+10) and 54 kcal/mol (+11) are similar to the E_a's measured for the nonspecific $(scFv + \mathbf{1}_{ns})^{n+}$ ions (Table 3.1), indicating that the binding affinity of the nonspecific ligand in the gas phase is relatively insensitive to the structure of the scFv and CA proteins.

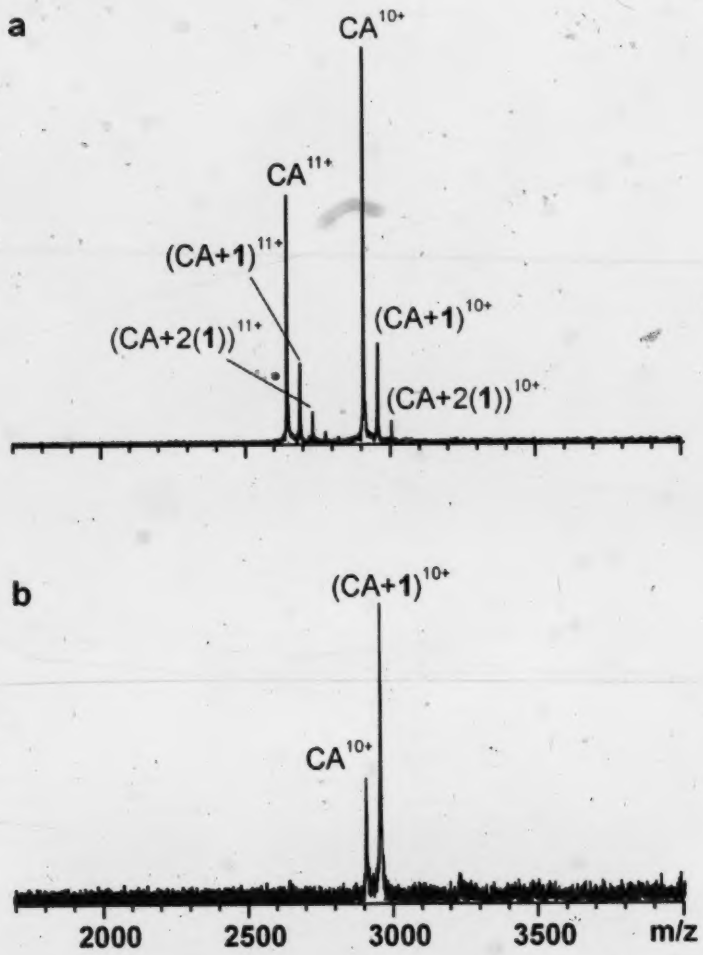


Figure 3.5 (a) NanoES mass spectrum obtained for an aqueous solution of CA (8 μM) and **1** (40 μM); (b) BIRD mass spectrum of the $(CA + \mathbf{1})^{10+}$ ion, reaction time 1.5 s, reaction temperature 165 °C.

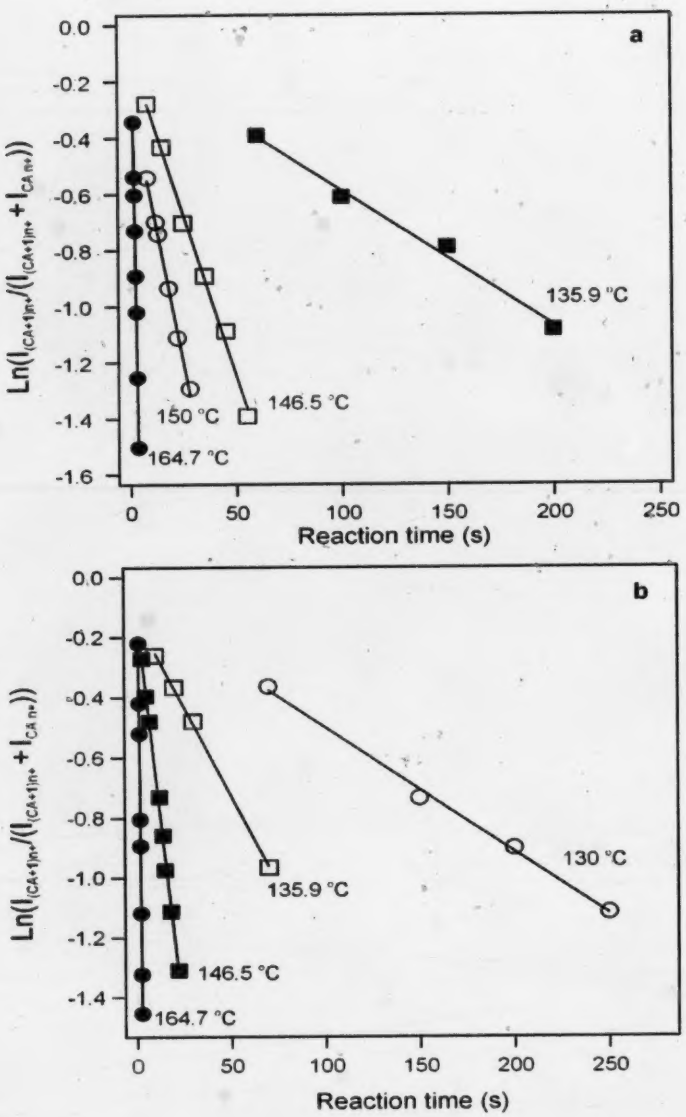


Figure 3.6 Plots of the natural logarithm of the normalized abundance of the protonated complex (a) $(CA + 1)^{10+}$ and (b) $(CA + 1)^{11+}$ versus reaction time at the temperatures indicated. Kinetic data fit to first-order kinetics.

3.4 Conclusions

In summary, we have determined for the first time kinetic and energetic data for the dissociation of protein-carbohydrate complexes produced from nonspecific interactions originating during the nanoES process. This study has revealed a number of unexpected results. (1) At the +10 and +11 charge states, the nonspecific $(scFv + \mathbf{1}_{ns})^{n+}$ ions are kinetically more stable than the corresponding $(scFv + \mathbf{1}_{sp})^{n+}$ ions. (2) The nonspecific complex at the +10 charge state is significantly energetically more stable than the corresponding specific complex. (3) The energetic stability of the nonspecific $(scFv + \mathbf{1}_{ns})^{n+}$ and $(CA + \mathbf{1})^{n+}$ ions are similar, suggesting that the nonspecific interactions are insensitive to the structure of protein. The findings of this study have implications for the determination of the affinity and stoichiometry of protein-ligand binding by nanoES-mass spectrometry. They also provide new insight into the intrinsic affinity and specificity of protein-carbohydrate binding.

3.5 Literature Cited

(1) (a) Hunter, C. L.; Mauk, A. G.; Douglas, D. J. *Biochemistry* **1997**, *36*, 1018-1025; (b) Wu, Q.; Gao, J.; Joseph-McCarthy, D.; Sigal, G. B.; Bruce, J. E.; Whitesides, G. M.; Smith, R. D. *J. Am. Chem. Soc.* **1997**, *119*, 1157-1158; (c) Rostom, A. A.; Tame, J. R. H.; Ladbury, J. E.; Robinson, C. V. *J. Mol. Biol.* **2000**, *296*, 269-279; (d) van der Kerk-van Hoof, A; Heck, A. J. R. *J. Mass Spectrom.* **1999**, *34*, 813-819.

(2) (a) Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 5902-5913; (b) Kitova, E. N.; Bundle, D. R.; Klassen, J.S. *J. Am. Chem. Soc.*

2002, *124*, 9340-9341; (c) Kitova, E. N.; Wang, W.; Bundle, D.R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 13980-13981.

(3) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S.-J.; MacKenzie, C.R.; Narang, S.A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. USA*, **1994**, *91*, 6423-6427.

(4) Kitova, E. N.; Klassen, J. S. manuscript in preparation.

(5) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945-4955.

(6) (a) Tholmann, D.; Tonner, D. S.; McMahon, T. B. *J. Phys. Chem.* **1994**, *98*, 2002-2004; (b) Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859-866.

(7) Felitsyn, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2001**, *73*, 4647-4661.

(8) Liu, D.; Wyttenbach, T.; Barran, P. E.; Bowers, M. T. *J. Am. Chem. Soc.* **2003**, *125*, 8458-8464.

Chapter 4

Nonspecific Protein-Carbohydrate Complexes Produced by Nanoelectrospray Ionization. Factors Influencing Their Formation and Stability*

4.1 Introduction

MS combined with ES or nanoES ionization is an established method for characterizing specific non-covalent biological interactions in solution and has been applied to a wide variety of complexes, including multiprotein assemblies, protein-ligand complexes, oligonucleotide duplexes and DNA- and RNA-drug complexes. The biochemical information available from ES-MS experiments is varied. ES-MS can detect specific complexes in buffered aqueous solutions and determine directly their binding stoichiometry [1]. ES-MS is also increasingly being used to measure the relative [2] and, in some cases, absolute [3] association constants (binding affinities). Recently, the application of variable-temperature nanoES-MS, for the quantification of association enthalpies and entropies for protein-ligand complexes [4] and for the determination of the thermal stability of protein assemblies in solution [5], was demonstrated. Combined with gas phase dissociation techniques, ES-MS is also being used to establish the composition [6] and, in some cases, the quaternary structure of biomolecular complexes [7]. Several review articles, which provide an overview of the many applications of ES-MS to biological complexes, have appeared [1f, 8].

While clearly a powerful bioanalytical tool, the ES-MS technique does have some

* A version of this chapter has been accepted by *Anal. Chem.*.

limitations. One arises from the tendency of biological molecules to associate nonspecifically with other biomolecules, small molecules or ions present in the ES solution. The resulting complexes may be sufficiently long-lived in the gas phase to be detected. The formation of non-covalent adducts with neutral or ionic species derived from the buffer or from impurities has been previously discussed and a variety of techniques for minimizing their formation have been reported. [9] In contrast, the tendency of biopolymers to associate into "nonspecific" complexes during the ES process and the nature of the stabilizing intermolecular interactions have not been systematically investigated.

The nonspecific association of biological molecules can complicate ES-MS experiments in a number of ways. Nonspecific complexes may be misidentified as "specific", *i.e.* false positives, and their presence may obscure the binding stoichiometry of specific complexes. It has also been shown that nonspecific binding between species that interact specifically in solution leads to binding affinities, determined by the direct ES-MS technique, which are artificially high. [3h, 10] There are two general strategies commonly employed with ES-MS to establish whether complexes observed in the mass spectra are specific or nonspecific in origin. These strategies have been discussed previously [8b, 11] and are only briefly described here. i) *Loss of specific interactions*. Modifying functional groups or residues that participate in specific interactions in solution will lead to changes in the relative abundance of specific complexes; complexes arising from nonspecific interactions are expected to be insensitive to such changes. ii) *Change in solution conditions*. Specific interactions are generally sensitive to solution conditions (*e.g.* pH, temperature, buffer) and changes will normally influence the relative

abundance of specific complexes, while the nonspecific interactions should be unaffected. It should be noted that the underlying assumption with both strategies, *i.e.* that response of the specific and nonspecific complexes to changes in structure or solution conditions will be very different, has not been widely tested and may not always be valid.

Where nonspecific complexes interfere with an ES-MS experiment, it may be desirable or necessary to eliminate them from the mass spectrum. In general, this can be achieved by working under conditions that minimize the formation of ES droplets that contain multiple analyte molecules (*e.g.* low concentration or small droplet volumes). However, this approach may not be feasible when studying weakly interacting species where high analyte concentrations may be necessary to detect the specific complexes. In such cases, it may be possible to reduce nonspecific interactions by working with near-equimolar solutions of protein and ligand, as shown in Chapter 2. It has also been proposed that the relative abundance of nonspecific complexes may be reduced by their selective dissociation in the gas phase [8b]. For such an approach to be feasible, the gaseous nonspecific complexes must be kinetically less stable than the specific complexes under the experimental conditions used. While this requirement may be satisfied in certain cases [12], it is not always valid [13, 14]. For example, a comparative study of the kinetic and energetic stability of a gaseous protein-trisaccharide complex originating from specific interactions in solution (and transferred to the gas phase by nanoES) and the corresponding nonspecific complex formed during the nanoES process was reported in Chapter 3. Using the BIRD technique [15] implemented with a FT-ICR/MS, time-resolved thermal dissociation experiments were performed on the

protonated specific and nonspecific complexes at two charge states, +10 and +11. This study revealed that the nonspecific complex was kinetically more stable than the specific complex, at both charge states. At the +11 charge state, the dissociation activation energies of the two complexes were indistinguishable, while at the +10 charge state, the nonspecific complex was energetically more stable. The origin of the greater stability of the nonspecific complex remains unclear although it was speculated that it might be due to the contribution of strong ionic intermolecular hydrogen bonds, interactions believed to be absent in the specific complex. Regardless of the nature of the nonspecific interactions, this study provides a clear example where gas phase dissociation reactions would not lead to the selective removal of nonspecific interactions.

A more detailed understanding of the factors responsible for stabilizing nonspecific complexes, as well as those that influence the formation of nonspecific complexes, would facilitate the design of ES-MS experiments which minimize nonspecific binding and, potentially, lead to new gas phase strategies to selectively remove ES-generated nonspecific complexes prior to MS detection. The structure and stability of nonspecific complexes of biologically-relevant molecules are also of fundamental interest as they reflect the affinity and specificity of biological interactions in the absence of solvent. Here, we report results from the first comprehensive investigation into the formation and stability of the nonspecific biological complexes. A series of nonspecific protein-carbohydrate complexes served as model systems for the study. Using nanoES-FT-ICR/MS, the influence of analyte concentration, the size and structure of the protein and carbohydrate, as well as charge state on the formation of nonspecific complexes was investigated. Thermal dissociation experiments, implemented

with the BIRD technique, were used to compare the kinetic stability of the nonspecific complexes and to suggest the nature of the stabilizing intermolecular interactions. The findings of this study, while relevant to the application of ES-MS for studying non-covalent interactions between proteins and small molecules or other biopolymers, also provide insight into the nanoES mechanisms responsible for the formation of gaseous ions of biological-relevant molecules.

4.2 Experimental

4.2.1 Proteins and Carbohydrates

CA (molecular weight ≡ MW 29 089 Da) and bovine ubiquitin, Ubq (MW 8 565 Da) were purchased from Sigma Canada and used without further purification. The carbohydrates, αAbe(2-O-CH_3-αMan) (**2**), αTal[αAbe]αMan (**3**), αAbe(2-O-CH_3-αMan)αGlcβGlc (**4**), were provided by D. R. Bundle (Univ. of Alberta), and D-Gal (**1**), D-Glc (**5**) and n-octyl-β-D-glucopyranoside (**6**) were purchased from Sigma Canada. The structures of these carbohydrates are shown in Figure 4.1.

4.2.2 Mass spectrometry

Gaseous complexes were produced by nanoES performed on aqueous solutions of protein (6 or 10 μM), carbohydrate (23 μM to 900 μM) and ammonium acetate (1 mM). The nanoES tips were constructed from aluminosilicate capillaries (1mm o.d., 0.68 mm i.d.), pulled to approximately 4 - 7 μm o.d. at one end using a P-2000 micropipette puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution was established by

applying a voltage of ± 800 V to a platinum wire inserted inside the glass tip. The solution flow rate was typically 20 to 50 nL/min. The droplets and gaseous ions produced by nanoES were introduced into the vacuum chamber of a modified ApexII 47e FT-ICR/MS (Bruker, Billerica MA) through a stainless steel capillary (i.d. 0.43 mm) maintained at an external temperature of 66 °C. [3h] The ion/gas jet sampled by the capillary (± 48 - 52 V) was transmitted through a skimmer (± 0 - 2 V) and stored in a hexapole. To minimize the extent of ion-source dissociation, a hexapole accumulation time of 1 s was used for complexes of CA and Ubq with the carbohydrates **1**, **2**, **5** and **6**. For all the remaining complexes, an accumulation time of 3 s was used. Ions were ejected from the hexapole and accelerated (± 2700 V) into a 4.7 tesla superconducting magnet, decelerated and introduced into the ion cell. The trapping plates of the cell were maintained at a constant potential of ± 1.4 – 1.8 V throughout the experiment. The typical base pressure for the instrument was ~ $5x10^{-10}$ mbar. The temperature of the ion cell was controlled with two external flexible heating blankets placed around the vacuum tube in the vicinity of the cell. [16]

Data acquisition was controlled by an SGI R5000 computer running the Bruker Daltonics XMASS software, version 5.0. Mass spectra were obtained using standard experimental sequences with chirp broadband excitation. Isolation of the reactant ions for the BIRD experiments was achieved using single rf frequency excitation. The isolated ions were stored inside the heated cell for varying reaction times prior to excitation and detection. The time-domain spectra, consisting of the sum of 30 transients containing 128 K data points per transient, were subjected to one zero-fill prior to Fourier-transformation.

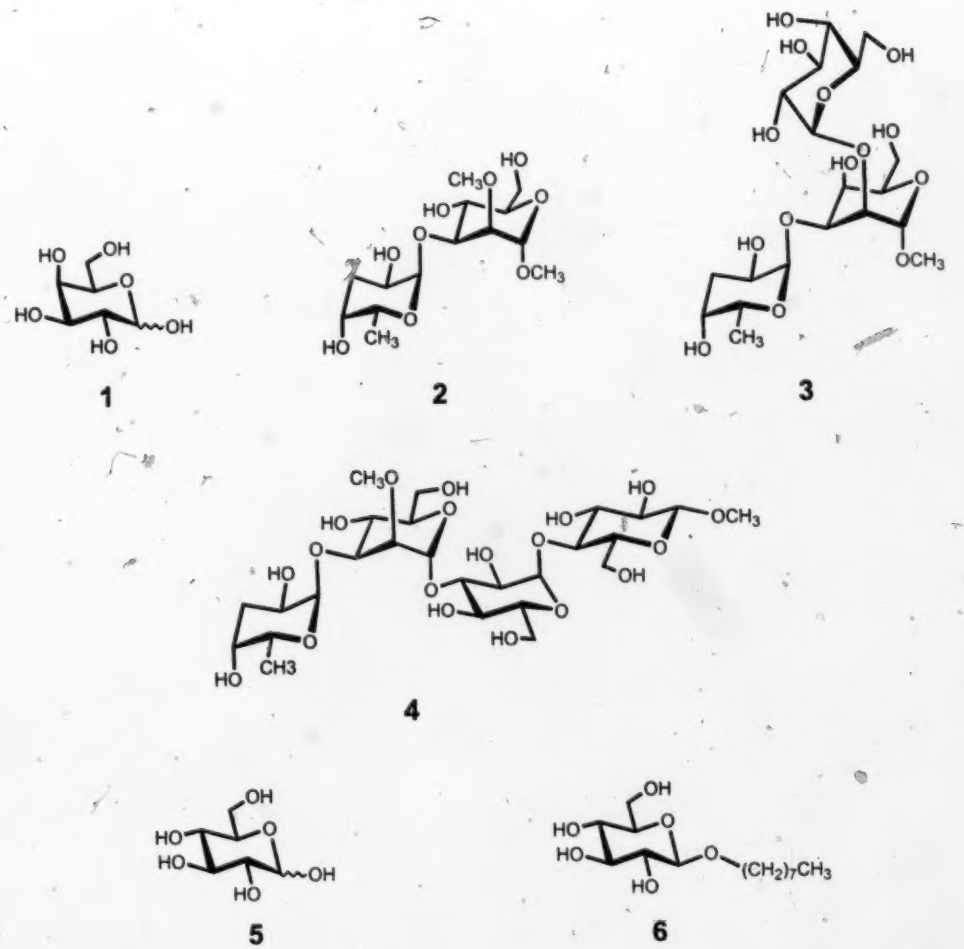


Figure 4.1 Structures of the carbohydrates: D-Gal (**1**), αAbe(2-O-CH_3-αMan) (**2**), αTal[αAbe]αMan (**3**), αAbe(2-O-CH_3-αMan)αGlcβGlc (**4**), D-Glc (**5**) and n-octyl-β-D-glucopyranoside (**6**).

4.3 Results and Discussion

4.3.1 Formation and Stabilization of Nonspecific Protein-Carbohydrate Complexes

To evaluate the propensity for proteins and carbohydrates to engage in nonspecific interactions during the nanoES process, "titration" experiments were performed on solutions containing either carbonic anhydrase (CA) or ubiquitin (Ubq) at a fixed concentration and one of the carbohydrate ligands, L = **1** – **6** (Figure 4.1), at varying concentrations. A comparison of the results provides insight into the influence of the nature of the protein and its charge state and of the carbohydrate on the formation of stable nonspecific complexes. Time-resolved thermal dissociation experiments were also performed on a number of complexes to evaluate their relative kinetic stability and to determine whether gas phase dissociation processes influence the relative abundance of nonspecific complexes measured by nanoES-MS.

4.3.1.1 Influence of Carbohydrate Structure: Size and Hydrophobicity

To assess the influence of ligand size on the formation/detection of nonspecific complexes, titration experiments were performed on solutions of CA, a 29 kDa Zn(II)-containing metalloenzyme, and one of the carbohydrates: D-Gal (**1**), αAbe(2-O-CH_3-αMan) (**2**), αTal[αAbe]αMan (**3**) or αAbe(2-O-CH_3-αMan)αGlcβGlc (**4**). These carbohydrates were selected because of the availability of kinetic data, determined by BIRD over a range of temperatures, for the dissociation of the corresponding 1:1 CA-carbohydrate complexes. [17] Shown in Figure 4.2a are nanoES mass spectra obtained in positive ion mode from aqueous solutions containing 6 μM CA and 23 μM L. The

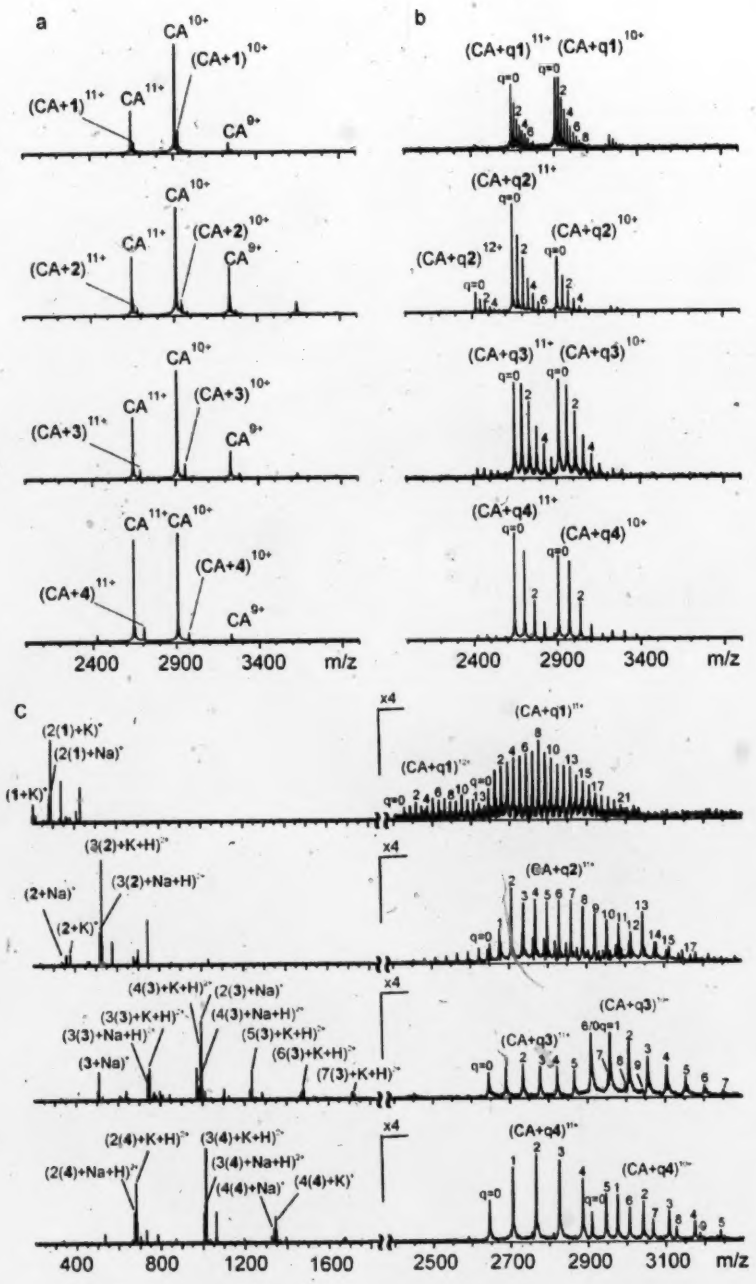


Figure 4.2 NanoES mass spectra obtained in positive ion mode for aqueous solutions containing CA (6 μM) and one of carbohydrates (L= **1** - **4**) at (a) 23 μM, (b) 140 μM and (c) 900 μM, with 1 mM ammonium acetate.

dominant ions observed in the mass spectra correspond to the protonated protein ions, $(CA + nH)^{n+} \equiv CA^{n+}$ at charge states of n = 9 – 12 with protonated 1:1 protein-carbohydrate complexes, $(CA + L + nH)^{n+} \equiv (CA + L)^{n+}$ with n = 9 – 12, also present in low abundance. CA has no known specific binding sites for any of these carbohydrates in solution and the 1:1 complexes observed in the mass spectra must result from nonspecific intermolecular interactions during the nanoES process. At much higher concentrations of L, ions of CA bound to multiple carbohydrate ligands are also observed, *i.e.* $(CA + qL)^{n+}$ where $q \geq 1$, and the breadth of the distribution of q for a given charge state increases with concentration of L (Figure 4.2b). This situation persists up to carbohydrate concentrations of ~ 1 mM, at which point the CA^{n+} and $(CA + qL)^{n+}$ ions are largely suppressed by ligand ions, $(pL + M)^{+}$ and $(pL + M + H)^{2+}$ where M = Na or K and p = 1 – 4 (Figure 4.2c). Somewhat larger clusters, with p as high as 7, are observed for the **3**.

A comparison of the results of the titration experiments performed with ligands **1** – **4** reveals a number of interesting features. First, the fraction of protein engaged in nonspecific interactions with one or more molecules of L increases with increasing ligand concentration, reaching a value of 0.95 at [L] = 900 μM (Figure 4.3). Secondly, the maximum number of bound carbohydrates, q_{max}, also increases systematically with [L], although, the number of bound carbohydrates is substantially less than would be predicted based on the initial ratio of carbohydrate to protein concentrations in solution. Interestingly, unbound CA^{n+} ions are observed over the entire range of concentrations investigated. Thirdly, at high carbohydrate concentrations, the distributions of q exhibit a striking dependence on the nature of L with a significantly broader distribution of q observed for the mono- (**1**) or disaccharide (**2**) compared to the solutions of the tri- (**3**) or

tetrasaccharide (**4**). For example, at a carbohydrate concentration of 900 μM, the values of q_{max} are 22 (**1**), 18 (**2**) and 9 (**3**, **4**), Figure 4.4. Finally, for ligands **1** and **2**, the magnitude of q_{max} observed at the highest ligand concentrations investigated exceeds the charge state of the complex, a result that suggests that the nonspecific association does not arise simply from charge solvation. Below, we provide explanations for the distribution of bound carbohydrates and the influence of concentration and carbohydrate structure thereon, in terms of the ES mechanism and the gas phase stability of the nonspecific complexes. In sections 4.3.1.2 and 4.3.1.3 we examine in more detail the role of charge on the formation and stabilization of the complexes.

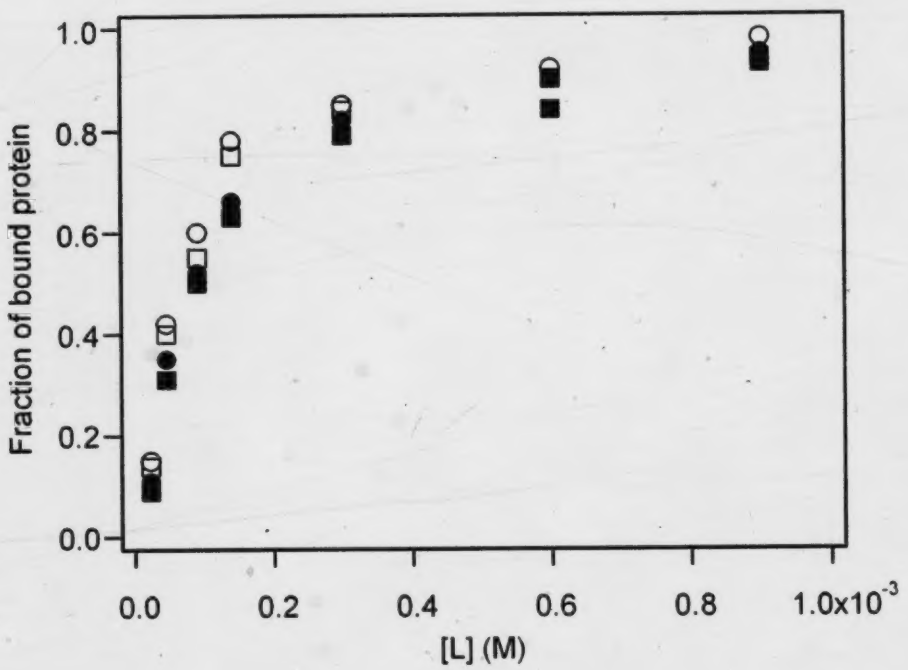


Figure 4.3 Plot of the fraction of gaseous CA ions engaged in nonspecific binding with L versus [L] measured for solutions containing 6 μM CA and L = **1** (○), **2** (●), **3** (□) **4**, (■).(Fraction of bound protein = 1- ($I_{\Sigma(CA^{n+}/n)}/I_{\Sigma(CA+qL)^{n+}/n}$)).

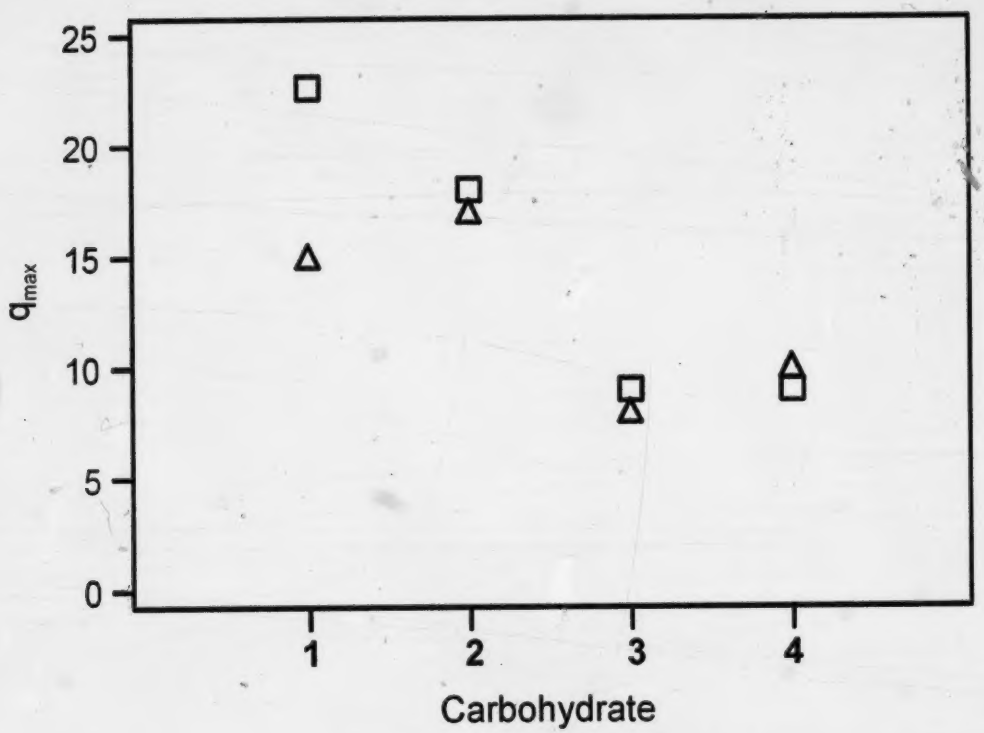


Figure 4.4 Plot of the maximum number of bound carbohydrate ligands (q_{max}) measured from aqueous solutions containing 6 μM protein (CA, □, and Ubq, Δ) and 900 μM L versus the number of saccharide residues: monosaccharide **(1)**, disaccharide **(2)**, trisaccharide **(3)** and tetrasaccharide **(4)**.

The increasing fraction of protein engaged in nonspecific interactions, as well as the increasing number of bound carbohydrates observed with increasing concentration is consistent with the formation of the gaseous nonspecific complexes via the charge residue ES model (CRM) [18]. CRM is the mechanism widely believed to be responsible for the formation of gaseous macromolecular ions, such as proteins and protein complexes, in ES. According to this model, the initial (parent) ES droplets undergo solvent evaporation until they come close to the Rayleigh limit, at which point they undergo fission, releasing several small multiply charged nanodroplets (offspring

droplets) containing none, one or multiple molecules of analyte. The parent droplets may repeat the evaporation/fission process many times, producing a large number of offspring droplet "litters" [19]. The details of the droplet fission processes in ES, and in particular for nanoES, are not fully known. However, based on the work of Gomez and Tang [20], it has been suggested that ~2% of the mass (volume) and ~20% of charge is transferred from the parent to the offspring droplets during fission. Solvent evaporation from the nanodroplets ultimately yields multiply charged gaseous ions. If the nanodroplets contain two or more analyte molecules, nonspecific intermolecular interactions will necessarily occur as the droplets evaporate to dryness. Because the probability of a droplet containing more than one analyte molecule increases with analyte concentration, the occurrence of nonspecific binding is also expected to increase with analyte concentration, which is consistent with the present experimental results.

The number of carbohydrates bound to the proteins is significantly smaller than the number that would be predicted based on the concentration ratio of carbohydrate and protein in solution ($[L]_o/[P]_o$). The fraction of carbohydrates that engage in long-lived nonspecific interactions with CA is more clearly seen in a comparison of the average number of carbohydrates expected to be bound to each protein, $q_{ave,calc}$, based on the ratio ($q_{ave,calc} = [L]_o/[P]_o$), compared with the average (weighted) number of bound ligands observed experimentally, $q_{ave,exp}$. The magnitude of $q_{ave,exp}$ was calculated using eq. 4.1:

$$q_{ave,exp} = \sum_q f_q q \qquad (4.1)$$

where f_q is the fractional abundance of complex ions with q bound ligands. Shown in Figure 4.5 is a plot of the ratio, $q_{ave,exp}/q_{ave,calc}$ versus [L] determined for **1** - **4**. For all four carbohydrates, the ratio initially increases with [L], reaching a maximum value of 0.04 -

0.1 at concentrations of 100 – 300 μM, then decreases with increasing concentration. It can also be seen that, at a given concentration, the magnitude of the ratio $q_{ave,exp}/q_{ave,calc}$ is sensitive to the nature of the carbohydrate, with the smallest molecules exhibiting the largest values. In other words, the smaller carbohydrates **1** and **2** form stable, nonspecific complexes more efficiently than do the larger carbohydrates **3** and **4**.

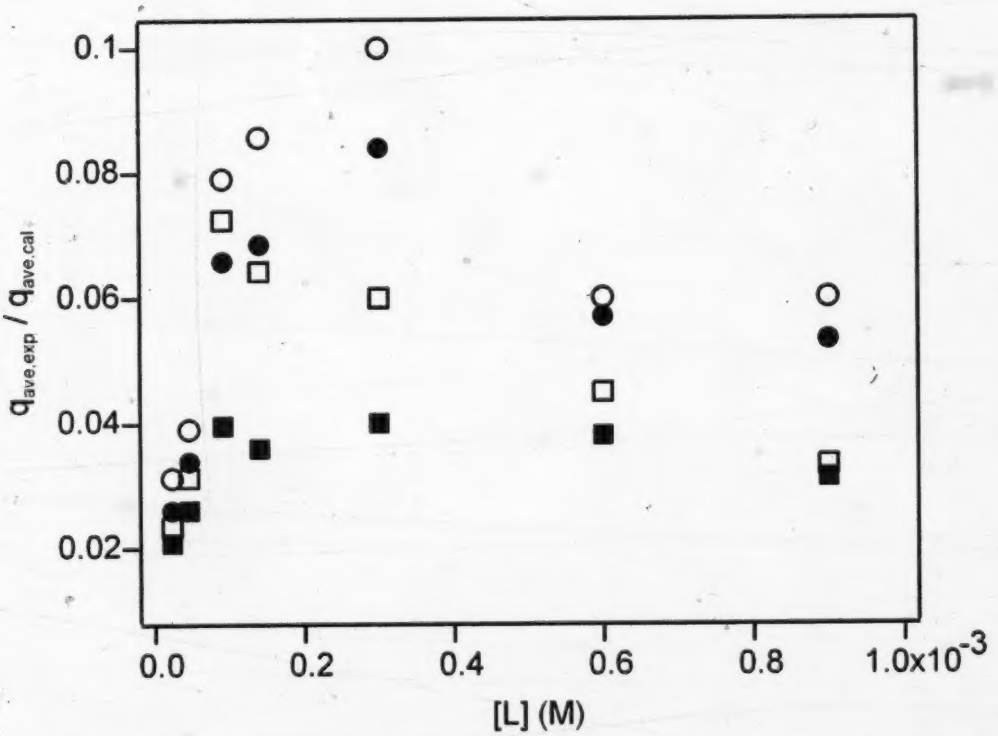


Figure 4.5 Plot of ratio $q_{ave,exp}/q_{ave,cal}$ versus [L] for solutions of 6 μM CA and L = **1** (○), **2** (●), **3** (□) **4**, (■). $q_{ave,exp}$ corresponds to the average number of bound carbohydrate ligands observed experimentally and $q_{ave,cal}$ is the predicted number based on the concentration ratio [L]/[P].

The general inefficiency of the nonspecific binding process (≤ 10%) and features of experimental distributions of q can be explained, at least in part, within the framework

of CRM. Assuming that the protein and carbohydrate molecules do not interact with themselves or each other in bulk solution or the parent droplets, and ignoring the influence of surface activity on the transfer of analyte to the offspring droplets, the protein and carbohydrate molecules will be randomly distributed between the offspring droplets. The average number of molecules of each analyte transferred to the fission droplets, λ, is related to the concentration of analyte in the parent droplets immediately prior to fission, $[X]_i$, and the volume of the offspring droplets, V_o, eq. 4.2:

$$\lambda = [X]_i V_o N_A \tag{4.2}$$

where N_A is Avogadro's number. The distribution of analyte between the offspring droplets (of a particular volume) is expected to follow Poisson statistics. The Poisson distribution of analyte in the offspring droplets is given by eq. 4.3:

$$P_x = (\lambda^x e^{-\lambda})/x! \tag{4.3}$$

where P_x is the probability of finding x analyte molecules in an offspring droplet. Although the experimental observation in the present work is the distribution of carbohydrates bound nonspecifically to protein, we assume that the distribution of carbohydrates in all offspring droplets will approximate the distribution of carbohydrate molecules in droplets containing a single protein. At low to moderate carbohydrate concentrations ($[L]_o \leq 300$ μM), a reasonably good fit between the experimentally observed distribution of q and the distribution calculated for a single value of λ is obtained. For example, the distributions obtained for solutions of CA with **1** and **4** at a concentration of 23 μM can be described by a Poisson distribution with a λ of 0.2 and 0.1, respectively (Figure 4.6a, b), while at 140 μM, the corresponding λ values are 1.3 and 0.9 (Figure 4.6c, d). These results, on their own, suggest that the gaseous nonspecific complexes are produced from a single set of offspring droplets or, if multiple generations

of offspring droplets are involved, they are characterized by similar $[X]_i V_o$ terms. At higher carbohydrate concentrations, however, the experimental distributions are too broad to be described by a single Poisson process. This is especially evident in the case of the smaller carbohydrate (Compare Figure 4.6e, f). Because the nonspecific complexes may originate from multiple generations of offspring droplets and the distribution may be influenced by additional ES and gas phase processes, *vide infra*, it is not possible to draw from the observed distributions any firm conclusions regarding the composition and size of the offspring droplets. Nevertheless, we felt that it was worthwhile to compare the experimental results with values of λ estimated using a procedure suggested by Kebarle and coworkers for assessing ES droplet histories.[19] Assuming an initial parent droplet radius in the range of 75 - 100 nm (which is reasonable given the solution flow rates used in the present work) [3h] and accounting for the influence of solvent evaporation on analyte concentration prior to fission [19], the values of λ for first generation fission droplets are estimated to be 0.1 – 0.2 (23 μM), 0.7 – 1.4 (140 μM) and 3.8 – 9.0 (900 μM). The estimated values of λ are in quite good agreement with the values found from the experimental distributions, at least at the lower carbohydrate concentrations investigated. This analysis suggests that, to a first approximation, the distribution of gaseous nonspecific complexes observed experimentally reflects the random distribution of carbohydrates between the offspring droplets. It follows then that it is the volume of the offspring droplets produced during fission that, to a large extent, determines the efficiency (or inefficiency) of the nonspecific binding process in nanoES. However, the

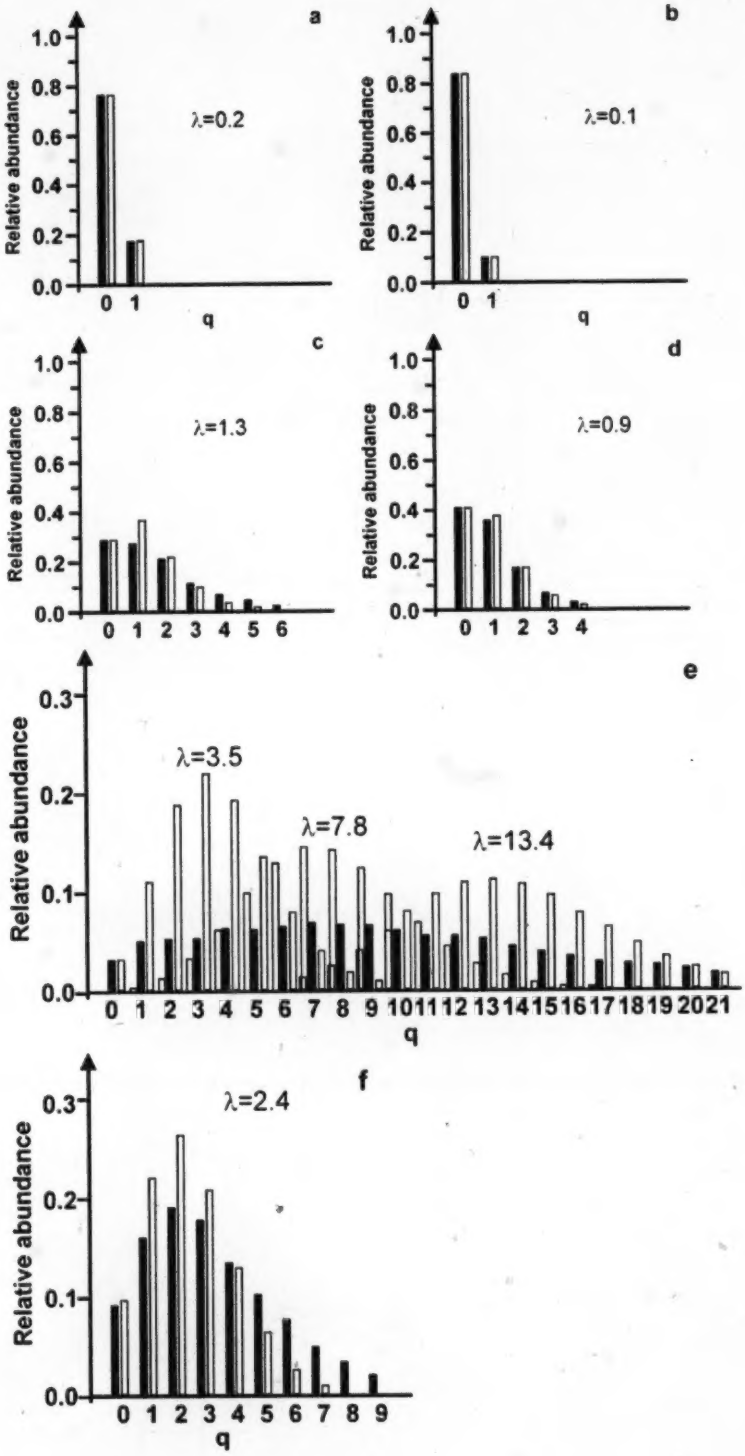

a
λ=0.2
Relative abundance
q
b
λ=0.1
c
λ=1.3
d
λ=0.9
e
λ=3.5
λ=7.8
λ=13.4
f
λ=2.4

Figure 4.6 Comparison of experimentally determined distributions (solid bars) of bound carbohydrates in the (CA + qL) ions, formed by nanoES from aqueous solutions containing CA (6 μM) and carbohydrate (L) (a) L = **1** at 23 μM, (b) L = **4** at 23 μM, (c) L = **1** at 140 μM, (d) L = **4** at 140 μM, (e) L = **1** at 900 μM and (f) L = **4** at 900 μM, with 1 mM ammonium acetate, and distributions calculated for a given λ using the Poisson formula, eq. 4.3, (open and shaded bars).

nonlinear dependence of λ (and $q_{ave,exp}/q_{ave,calc}$) on $[L]_o$, as well as its dependence on the structure of the carbohydrate can not be explained simply in terms of the random distribution of carbohydrates in the offspring droplets. Additional processes must also influence the efficiency of nonspecific binding and the distribution of nonspecific complexes observed experimentally. Below we describe the effects of ion evaporation, cluster formation and gas phase dissociation on the formation and detection of the nonspecific complexes.

Due to their small size, the carbohydrate molecules are expected to be able to evaporate, as ions, from the charged offspring droplets. As a result, the number of carbohydrate molecules contained within the offspring droplets prior to the formation of the gaseous complexes may be less than the number initially transferred during fission. According to the ion evaporation model (IEM) [21], the efficiency of ion evaporation from the surface of a charged droplet is related to the solvation free energy of the ion. Ions which are not strongly solvated, *i.e.* which have a high surface activity, are more rapidly lost through evaporation.[22] Differences in surface activity may arise from differences in chemical surface activity as well as charge-induced surface activity. In a recent study by Karas and coworkers [23], it was shown that carbohydrates which contain a hydrophobic moiety exhibit greater ionization efficiencies than unmodified carbohydrates. To demonstrate the influence of ion evaporation on the efficiency of nonspecific protein-carbohydrate complex formation, measurements were performed on solutions of CA with D-glucose (**5**) and with n-octyl-β-D-glycopyranoside (**6**), which is expected to be significantly more surface active than **5**. Shown in Figures 4.7a,b are mass spectra obtained from solutions of 6 μM CA and 720 μM **5** or **6**. It can be seen that the

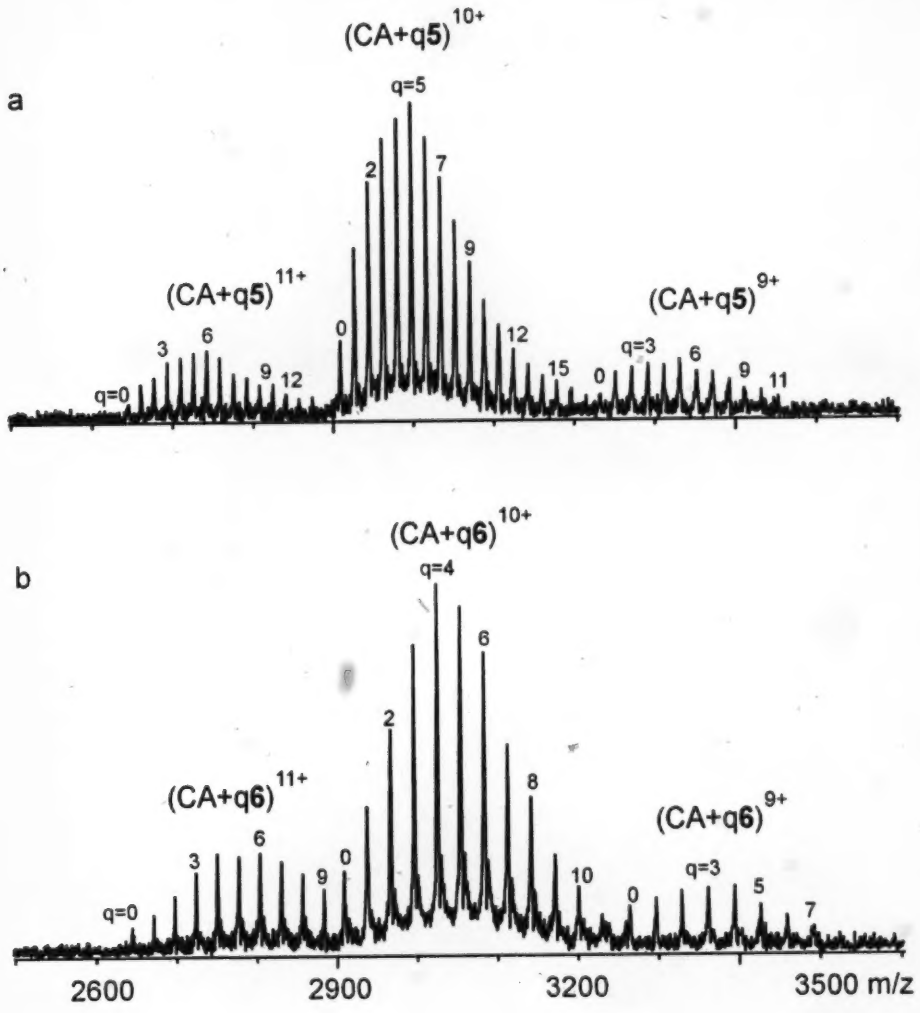


Figure 4.7 NanoES mass spectra obtained in positive ion mode for aqueous solutions of CA (6 μM) and 720 μM (a) **5** or (b) **6**, with 1 mM ammonium acetate.

distribution of q for **6** is somewhat narrower than for **5**, with a q_{max} of ~11 for **6**, compared to ~16 for **5**. Thermal dissociation experiments performed at 130 °C, a reaction temperature that is close to the estimated temperature of the ions within the ion source,

vide infra, reveal that the complexes of **6** are kinetically more stable than those of **5** (see Figure 4.8).

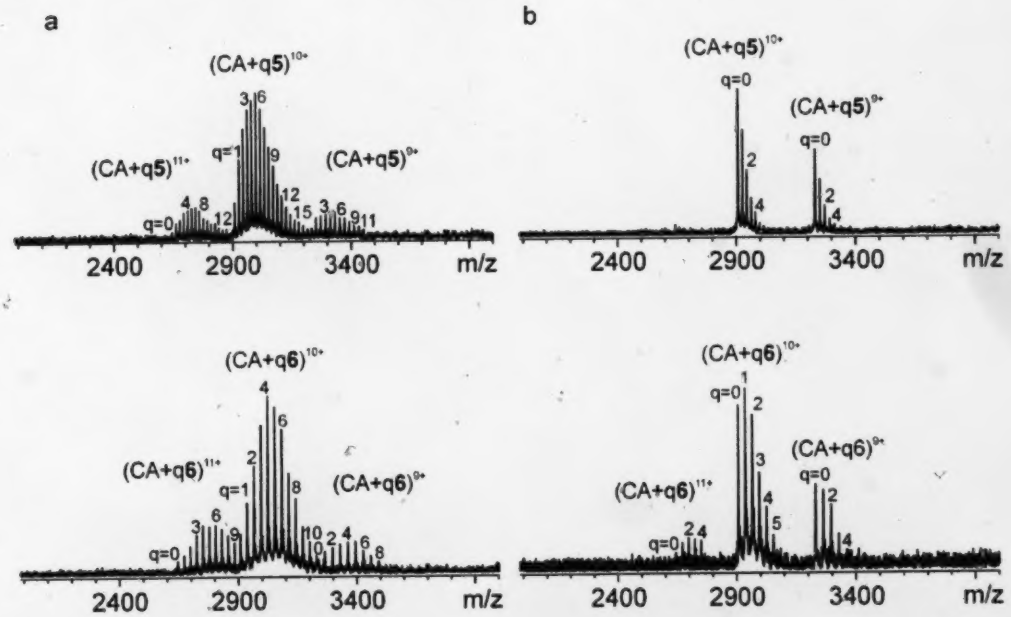


Figure 4.8 (a) NanoES mass spectra obtained from the solutions of 6 μM CA and 720 μM **5** or **6**. (b) Corresponding BIRD mass spectra obtained at a cell temperature of 130 °C and a reaction time of 1 s.

Consequently, the narrower distribution observed for **6** does not arise from in-source dissociation but, instead, must reflect the influence of carbohydrate surface activity on the formation of nonspecific complexes, with a lower surface activity promoting association. The reduction in nonspecific binding observed for **6** compared to **5** is consistent with a greater loss of **6** from the offspring droplets by ion evaporation. This result further suggests that the influence of ligand size on q_{max} (and λ) could result from the preferential evaporation of the larger carbohydrates. Additional support for this conclusion can be found in a comparison of the nanoES ionization efficiencies of the four

ligands, **1** – **4**. Shown in Figure 4.9 is a nanoES mass spectrum obtained from an aqueous equimolar (4 μM) solution of **1**, **2**, **3** and **4**. The major species observed are the $(L + Na)^+$ and $(L + K)^+$ ions. It can be seen that their relative abundance increases with the size of the carbohydrate. Assuming the gaseous carbohydrate ions are produced by IEM and their relative abundance reflects their ionization efficiencies, this result indicates that the ionization efficiency of the carbohydrates increases with their size. This trend in ionization efficiency presumably reflects the trend in ion solvation energy, which is expected to decrease with increasing size of L due to more efficient solvation of the charge by the carbohydrate molecule.

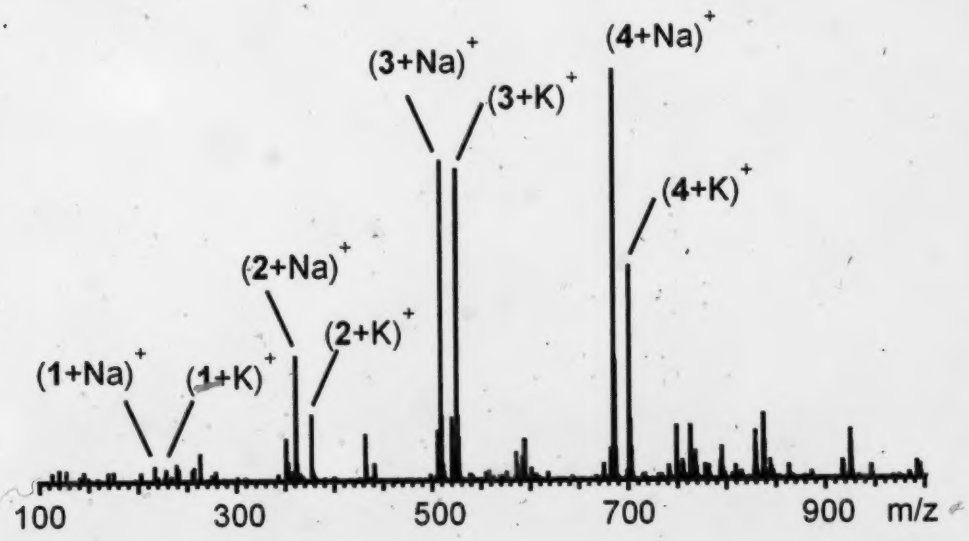


Figure 4.9 NanoES mass spectrum obtained from an aqueous equimolar (4 μM) solution of **1**, **2**, **3** and **4**, with 1 mM ammonium acetate.

The tendency of the carbohydrates to self-associate into nonspecific clusters during the nanoES process may also influence the extent of nonspecific binding. In particular, cluster formation could account for the decrease in the binding efficiency

observed at concentrations above 300 μM (Figure 4.4). In Figure 4.2c it can be seen that, at high carbohydrate concentrations, charged carbohydrate clusters of the type, $(pL + M)^+$ and $(pL + M + H)^{2+}$, are observed. The phenomenon of cluster formation has been previously demonstrated for other neutral species, including amino acids [24], peptides [25] and salts [26]. The ES mechanism(s) leading to cluster formation is not fully resolved, with evidence for the contribution of both CRM and IEM. [24a,26e-h] If CRM operates, cluster formation is expected to occur simultaneously and competitively with the formation of the nonspecific protein-carbohydrate complexes and, presumably, result in a reduction in the number of carbohydrates bound to the protein, Figure 4.10.

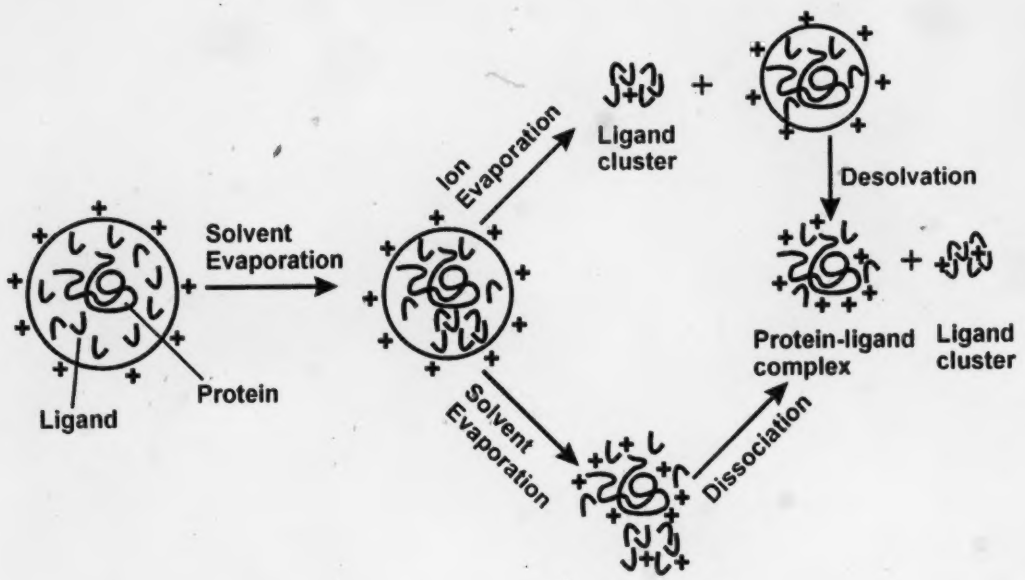


Figure 4.10 Formation of nonspecific protein-ligand complexes and ligand clusters as solvent evaporates to dryness during the nanoES process via IEM and CRM.

Alternatively, cluster formation by IEM prior to the formation of the protein-carbohydrate complexes would reduce the number of carbohydrate molecules available within the offspring droplets. Therefore, regardless of the mechanism responsible for their formation, the self-association of the carbohydrates during the nanoES process is expected to lead to a reduction in the abundance and extent of nonspecific protein-carbohydrate binding. It is also worth noting that the tendency of a carbohydrate to self-associate into clusters appears to be influenced by its size. It can be seen in Figure 4.2c that **3** and **4** tend to form larger clusters more so than **1** and **2**. The greater tendency of the larger carbohydrates to form clusters may explain their reduced efficiency at forming nonspecific complexes with the protein.

The dissociation of nonspecific protein-carbohydrate complexes in the gas phase prior to detection will also reduce the number of bound carbohydrates. The lifetime of a gaseous noncovalent complex is governed by the nature and strength of the intermolecular interactions and its internal energy. The internal energy of ions sampled into the mass spectrometer will depend on the sampling conditions (*e.g.* sampling capillary temperature, nozzle-skimmer potential, background pressure) as well as the size and charge state of the ion. As described previously, [3h, 27] the internal energy of the ions sampled into the ion source used for the present measurements is most significantly influenced by energetic collisions within the rf hexapole. Collisional heating of the ions during accumulation in this region can cause weakly-bonded noncovalent complexes to dissociate prior to detection. In a previous study, it was shown that the +9 to +11 ions of a 27 kDa protein-trisaccharide complex had an "effective" temperature of ~140 °C in the hexapole region. [3h] The $(CA + qL)^{n+}$ ions studied here, which are of comparable mass

and charge state, are expected to be similarly heated in the hexapole. To assess the tendency of the $(CA + L)^{n+}$ complexes to dissociate in the source region, the hexapole accumulation time was varied, between 1 s (which we find is the minimum accumulation time required to obtain spectra of reasonable S/N) and 3 s, for each complex. Illustrative nanoES spectra obtained at two different accumulation times (1 s and 3 s) for the aqueous solutions of CA and **1** or **4** are shown in Figure 4.11. It was found that the distribution of $(CA + q\mathbf{3})^{n+}$ and $(CA + q\mathbf{4})^{n+}$ ions was insensitive to accumulation time, while the distribution of the $(CA + q\mathbf{1})^{n+}$ and $(CA + q\mathbf{2})^{n+}$ ions became narrower (and q_{max} decreased) with increasing time. Consequently, all measurements reported for the mono- and disaccharides were obtained using an accumulation time of 1 s, which minimizes in-source dissociation.

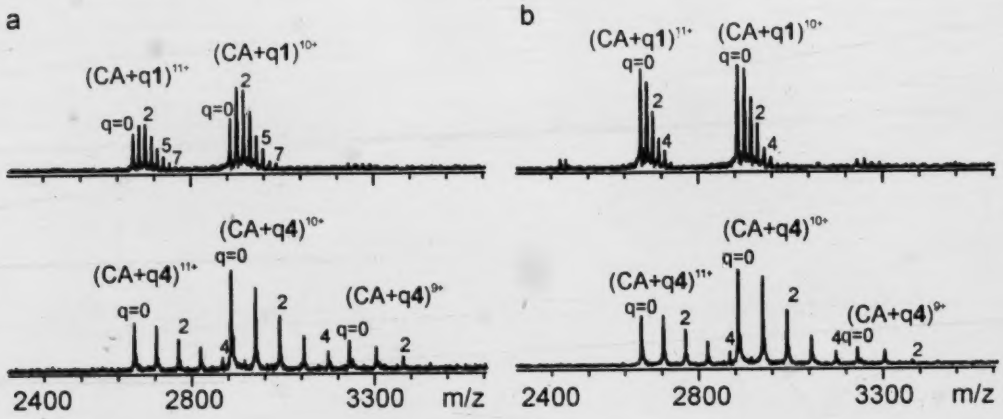


Figure 4.11 NanoES mass spectrum obtained from aqueous solutions containing 6 μM CA and 200 μM of **1** and **4**, with 1 mM ammonium acetate with an accumulation time of (a) 1 s and (b) 2 s in the hexapole.

With the exception of Arrhenius parameters determined by BIRD for the dissociation of two 1:1 nonspecific protein-trisaccharide complexes, including the (CA +

3)$^{n+}$ complex at n = 10, 11, [14] kinetic and energetic data have not been reported for gaseous nonspecific protein-ligand complexes. To evaluate the kinetic stability of the gaseous (CA + qL)$^{n+}$ ions, BIRD snapshot experiments were performed. In these experiments, no ion isolation step is performed prior to reaction. Instead, all of the (CA + qL)$^{n+}$ present in the ion cell are allowed to dissociate simultaneously. Shown in Figures 4.12a,b are nanoES spectra obtained from the aqueous solutions containing 6 μM CA and 0.3 mM **1**, **2**, **3** or **4** and the corresponding BIRD snapshot spectra obtained at a cell temperature of 160 °C and a reaction time of 1 s, respectively. Because the loss of carbohydrate does not cause any significant change in the charge state distribution, it can be concluded that the carbohydrate molecules are lost predominantly from the complex in their neutral form. It can be seen that the extent of dissociation of the (CA + qL)$^{n+}$ ions decreases with increasing ligand size. For example, BIRD of the (CA + q**1**)$^{n+}$ ions, where $q_{max} \approx 14$, resulted exclusively in the unbound CA^{n+} ions, while the distribution of the (CA + q**4**)$^{n+}$ ions, where $q_{max} \approx 6$, was largely unaffected. The BIRD results reveal that, at 160 °C, the kinetic stability of the gaseous (CA + qL)$^{n+}$ ions increases with the size of the ligand: **1** < **2** < **3** < **4**. The trend in stability is consistent with that found in a related study of the dissociation kinetics of the corresponding 1:1 (CA + L)$^{n+}$ ions at reaction temperatures ranging from 60 to 190 °C. [17] This trend in stability for the (CA + qL)$^{n+}$ ions likely reflects the tendency of the larger carbohydrates, which possess more oxygen groups (-OH, -OCH_3, -O-), to form more intermolecular hydrogen bonds which result in energetically and kinetically (at this temperature) more stable complexes. The results of the thermal dissociation experiments indicate that the decrease in nonspecific protein-carbohydrate binding with increasing size of the carbohydrate does not reflect differences

in gas phase stability of the resulting complexes. It should be noted, however, that the BIRD experiments only provide insight into the relative stability of complexes that survive long enough to reach the ion cell. It is possible that some of the nonspecific protein-carbohydrate complexes produced by nanoES have lifetimes much shorter than the sampling times used in the present measurements and that they dissociate in the source region.

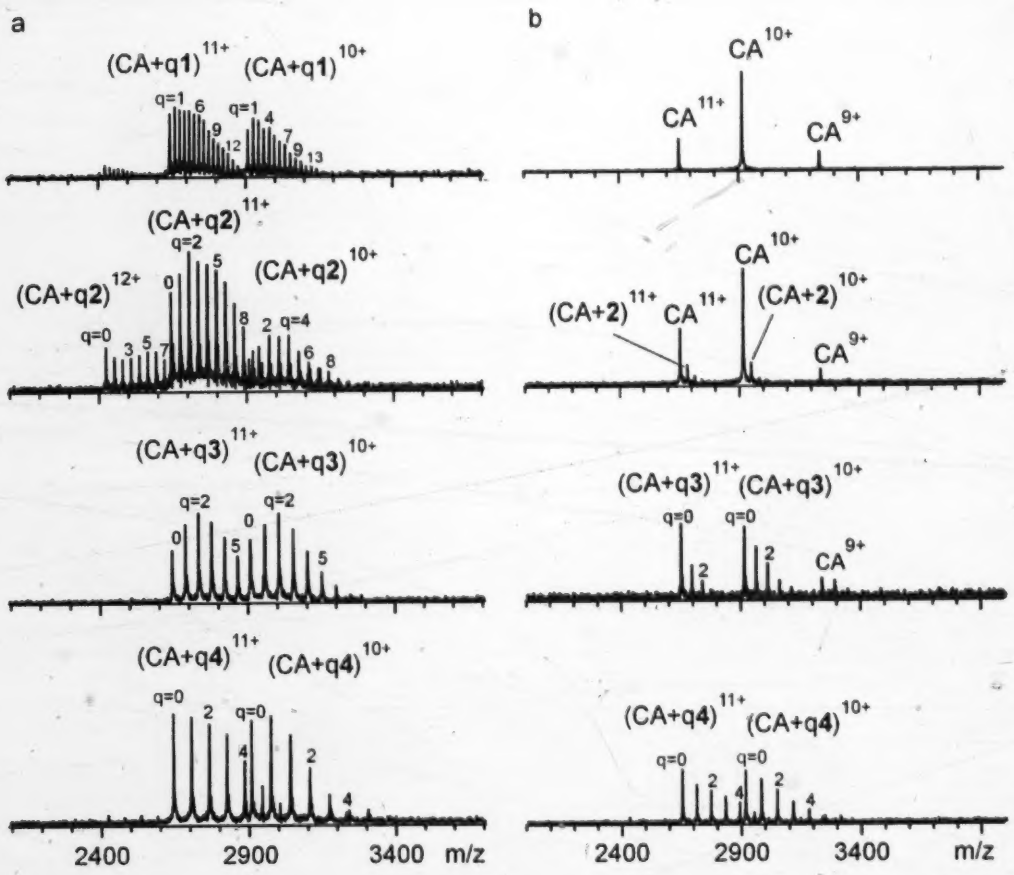


Figure 4.12 (a) NanoES mass spectra obtained from the solutions of 6 μM CA and 300 μM L (**1 - 4**). (b) BIRD mass spectra obtained at 160 °C and a reaction time of 1 s.

4.3.1.2 Influence of Protein Size and Structure

The results from the preceding section clearly demonstrate the influence of the carbohydrate structure on the formation and stability of nonspecific protein-carbohydrate complexes produced by nanoES, with small, hydrophilic carbohydrates more efficiently forming nonspecific complexes, albeit of lower gas phase stability, than larger or more hydrophobic carbohydrates. The nature of the protein, in particular the surface presented by the protein (area and structural elements) to the carbohydrates, might also influence the degree of binding and the stability of the nonspecific complexes. To establish whether the structure of the protein influenced the formation of the nonspecific complexes, titration experiments were performed with **1** – **4** and ubiquitin (Ubq) and the results compared with those obtained for CA. Ubq is an 8.5 kDa cytoplasmic protein with an extended β-sheet and an α-helix surrounding a hydrophobic core. [28] From an analysis of their crystal structures [29], the surface area of Ubq (4792 $Å^2$) is estimated to be less than one third that of CA (16264 $Å^2$). Surface area calculations were also performed for **1** – **4** for comparison: **1** (322 $Å^2$), **2** (537 $Å^2$), **3** (675 $Å^2$), **4** (903 $Å^2$). Shown in Figure 4.13 are nanoES mass spectra obtained from aqueous solutions containing Ubq (6 μM) and each of the carbohydrates, **1** – **4**, at 900 μM. The spectra contain abundant protonated $(Ubq + qL)^{n+}$ ions, with n = 5 and 6. From a comparison of mass spectra in Figures 4.2c and 11 it can be seen that, despite differences in the surface area and structure of CA and Ubq, and the charge states of the complexes, the distribution of q is similar for L = **2** – **4** (see Figure 4.4 for a comparison of q_{max} values). The similarity in the extent of nonspecific binding observed for **2** – **4** with CA and Ubq suggests that the size and structure of the protein are not determining factors in the formation of nonspecific

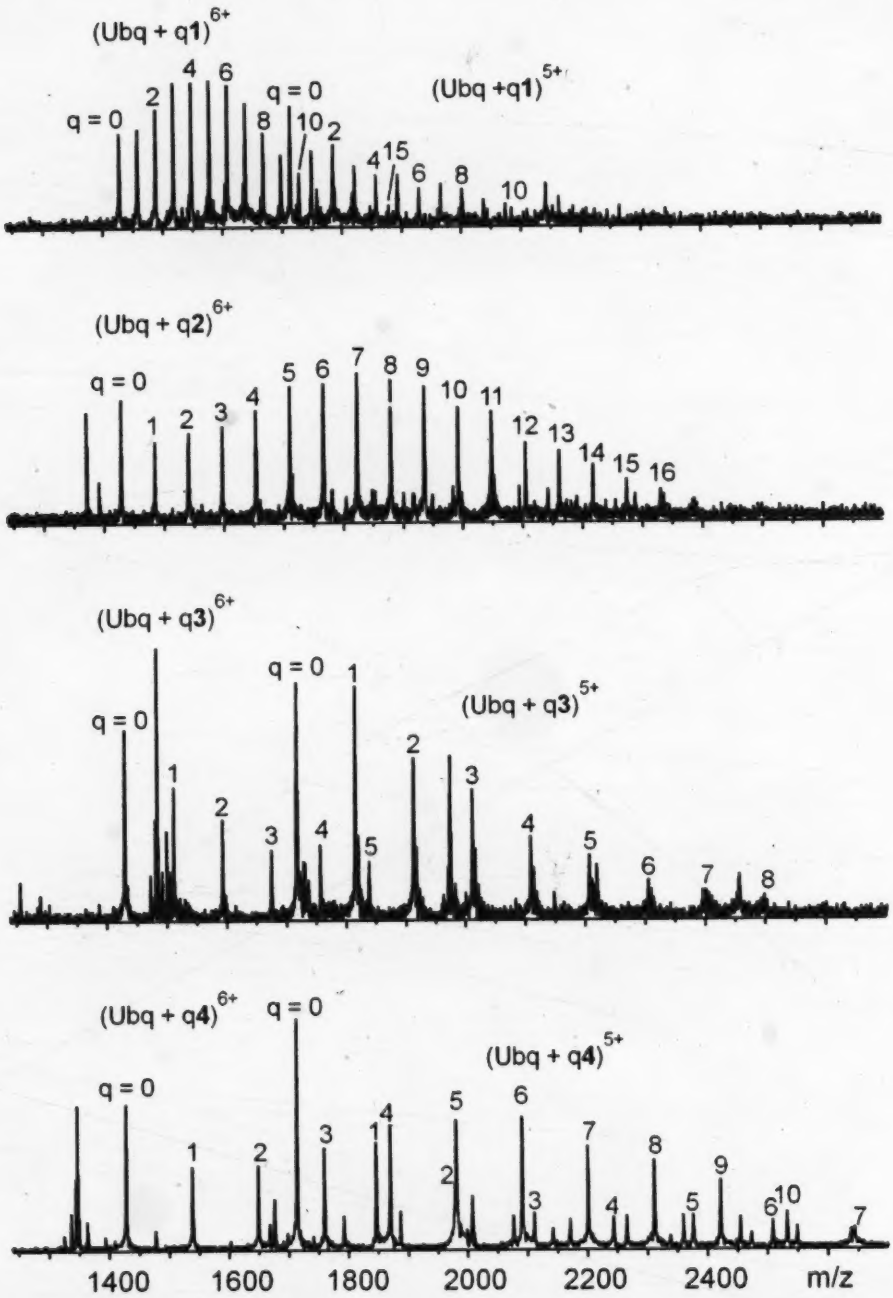


Figure 4.13 NanoES mass spectra obtained in positive ion mode for aqueous solutions containing Ubq (6 μM) and one of carbohydrates (L = **1** - **4**) at 900 μM.

complexes. It should be noted that, assuming each bound carbohydrate interacts directly with the protein and occupies an area on the protein equal to ~50% of its own surface area, the surface of the protein does not become saturated with bound carbohydrates even at the highest carbohydrate concentrations studied. In other words, for both CA and Ubq, the available surface area exceeds the area required for each carbohydrate to bind directly to the protein. It is possible that the surface area available with smaller proteins would influence the degree of carbohydrate binding.

For L = **1** (at 900 μM), a much narrower distribution of q is observed for Ubq (q_{max} ~15) compared with CA (q_{max} ~22) (Figure 4.4). This difference in the degree of binding could reflect differences in kinetic stability of the Ubq and CA complexes or differences in the efficiency of nonspecific binding during the nanoES. Support for the former explanation can be found in comparison of BIRD snapshot data obtained for the $(CA + q\mathbf{1})^{n+}$ and $(Ubq + q\mathbf{1})^{n+}$ ions. Shown in Figures 4.14a,b are mass spectra obtained for solutions of CA and Ubq with **1** and the corresponding BIRD spectra obtained at a temperature of 110 °C and a reaction time of 1 s, respectively. While the distribution of $(CA + q\mathbf{1})^{n+}$ ions are unchanged in the two spectra, the distribution for the $(Ubq + q\mathbf{1})^{n+}$ ions is significantly reduced in the BIRD spectrum, with most of the carbohydrates lost after a 1 s reaction. These results indicate that the $(Ubq + q\mathbf{1})^{n+}$ complexes are kinetically less stable than the $(CA + q\mathbf{1})^{n+}$ complexes at this temperature and, therefore, are expected to undergo dissociation in the source region of the mass spectrometer. It is likely that the greater kinetic stability of the CA complexes, compared to the Ubq complexes, is a reflection of their greater energetic stability. Although this remains to be

verified, it has been shown to be true for the 1:1 complexes of CA and Ubq with **3** at the charge states investigated here. [17]

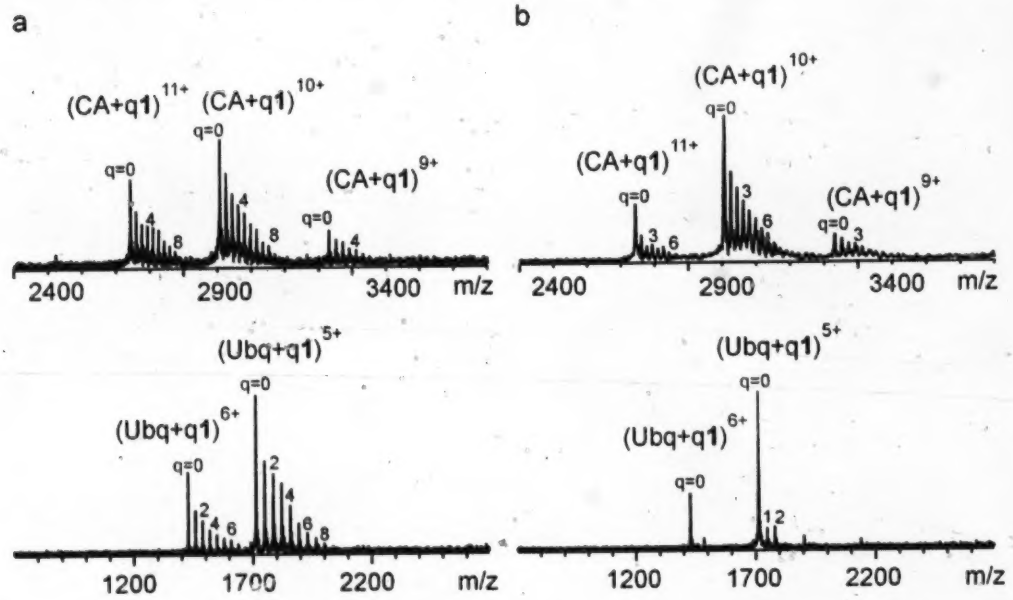


Figure 4.14 (a) NanoES mass spectra obtained from the solutions of 6 μM CA or Ubq and 100 μM **1**. (b) BIRD mass spectra obtained at 110 °C and a reaction time of 1 s.

4.3.1.3 Influence of Charge

The observation that the distribution of q is relatively insensitive to the charge state of the protonated complexes and that the q_{max} can, at high concentrations of L, exceed significantly the charge state of these complex, indicates that extent of nonspecific association during nanoES is not governed by the number of available charges. A comparison of the mass spectra obtained for solutions of Ubq and **2** in positive and negative ion mode further highlights the absence of a charge state effect on the

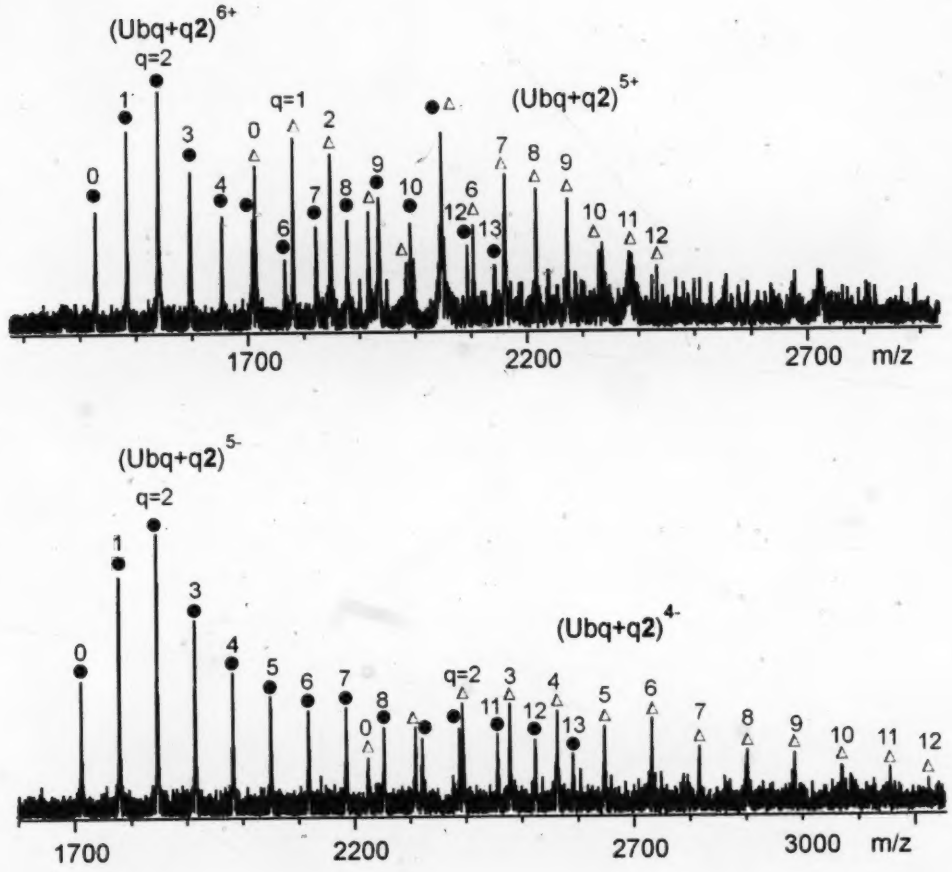


Figure 4.15 NanoES mass spectra obtained from aqueous solutions containing Ubq (10 μM) and 2 (400 μM) in (a) positive and (b) negative ion modes.

formation of the nonspecific complexes. Shown in Figure 4.15 are mass spectra obtained in positive and negative ion mode from a solution of Ubq (10 μM) and **2** (400 μM). It can be seen that the distributions of q measured for the +6 and +5 charge states are very similar to the distributions measured for the –5 and –4 charge states, indicating that the formation of the nonspecific complexes is insensitive to nature of the charge groups. Similar distributions, in positive and negative ion mode, were also observed for the

complexes of Ubq with the carbohydrates **1**, **3** and **4**. These results clearly demonstrate that these nonspecific protein-carbohydrate complexes do not originate solely from strong ionic intermolecular interactions (*i.e.* charge solvation) between the protein and carbohydrate; neutral interactions must also contribute significantly and, perhaps, predominantly.

Also considered in this study was the influence of charge on the kinetic stability of the complexes. For a given protein, the number and type of charge groups can, in principle, influence the stability of the noncovalent complexes in two ways. First, the stability of bound carbohydrates stabilized wholly or in part by ionic interactions will be sensitive to the number of charges present, the nature of the charge group and its chemical environment. Secondly, the three dimensional structure of multiply charged gaseous proteins are known to be sensitive to electrostatic effects. [30] Differences in structure may, in turn, lead to differences in the number and the strength of the intermolecular interactions. To investigate the influence of charge on the stability of the complexes, BIRD was performed on the $(CA + qL)^{n+}$ ions, where n = 9 – 11, and $(Ubq + qL)^{n+}$ ions, where n = 4 – 6, for L = **1**, **3**. For these measurements, isolation of the complexes at the desired charge state was performed prior to the reaction. Relatively high reaction temperatures were chosen to achieve a significant degree of dissociation at short reaction times and, thereby, maintain a high ion signal-to-noise ratio (S/N). Shown in Figures 4.16a,b are mass spectra obtained after isolation of the $(CA + q\mathbf{3})^{n+}$ ions, where q_{max} ranged from 3 - 5, at a given charge state and BIRD spectra measured at a temperature of 167 °C and a reaction time of 2s, respectively. A comparison of these spectra reveals that the fraction of CA involved in nonspecific interactions decreased by

40% for +9, 20% for +10 and 55% for +11 at the end of the 2 s reaction. Based on these results, the trend in kinetic stability at this temperature is: +10 > +9 > +11.

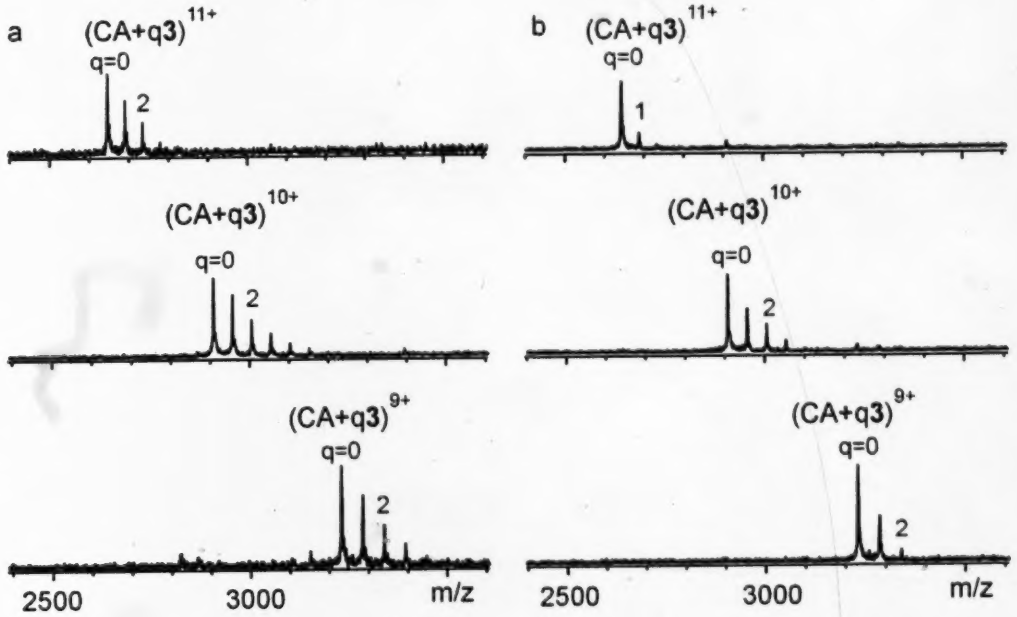


Figure 4.16 (a) Isolation mass spectra obtained from the solutions of 6 μM CA and 100 μM **3**; (b) BIRD mass spectra obtained at 167 °C and a reaction time of 2 s.

The same trend in reactivity was found for the (CA + q**1**)$^{n+}$ ions of the same charge states at 150 °C (see Figure 4.17a, b). Analysis of the BIRD results for the (Ubq + q**3**)$^{n+}$ ions obtained at a temperature of 160 °C and a reaction time of 1.5 s (Figures 4.18 a,b) reveals that the fraction of Ubq involved in nonspecific interactions decreased by 35% for +4, 20% for +5 and 30% for +6. Based on these results, the trend in kinetic stability at this temperature is: +5 > +6 ≈ +4. Similar results were also obtained for the (Ubq + q**1**)$^{n+}$ ions

of the same charge states at 100 °C. To assess the importance of the nature of the charge groups, specifically the influence of protonated versus deprotonated groups, on kinetic

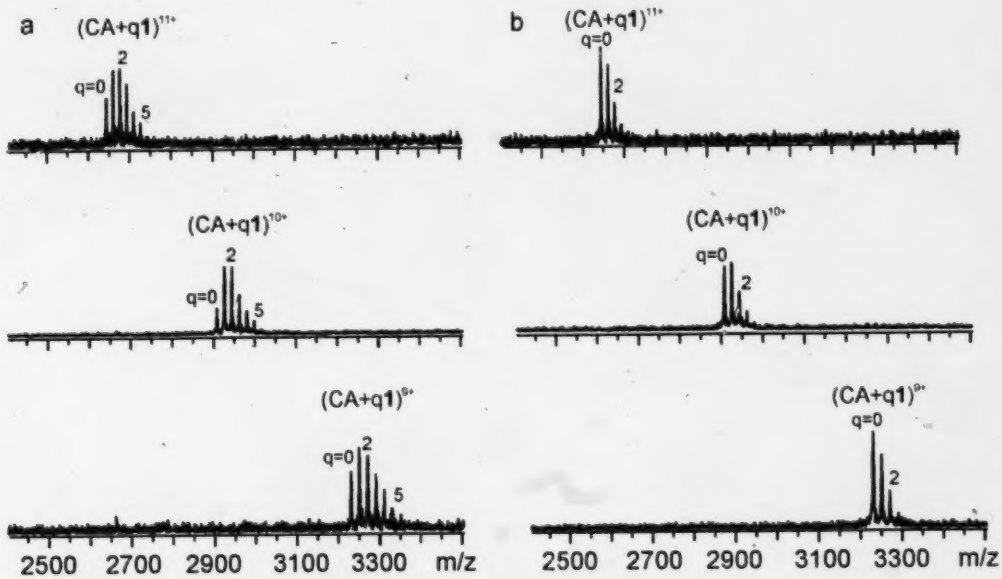


Figure 4.17 (a) Isolation mass spectra obtained from the solutions of 6 μM CA and 180 μM **1**; (b) BIRD mass spectra obtained at 150 °C and a reaction time of 1 s.

stability, BIRD snapshot experiments were performed on $(Ubq + qL)^{n+/z-}$ ions, where n = 4 – 6 and z = 4, 5, for L = **1 - 4**. Illustrative BIRD spectra obtained for the complexes of Ubq with **1 - 4** are shown in Figures 4.19 and 4.22, respectively. It is found that, for **1** and **2**, the protonated complexes are kinetically more stable than corresponding deprotonated complexes. A similar result is found for the complexes of **4**. However, it should be noted that the major dissociation pathway for the deprotonated complexes of **1 - 3** is the loss of neutral carbohydrate while for ligand **4** it is the loss of the deprotonated carbohydrate. Interestingly, for the complexes of **3** deprotonated complexes are found to be slightly more stable than the corresponding protonated complexes (Figure 4.21). Although far

from an exhaustive study, the BIRD measurements demonstrate that the number and nature of the charge groups do influence the kinetic stability of the nonspecific complexes, although the differences in stability tend to be relatively subtle. Furthermore, no simple relationship between stability and charge state could be identified.

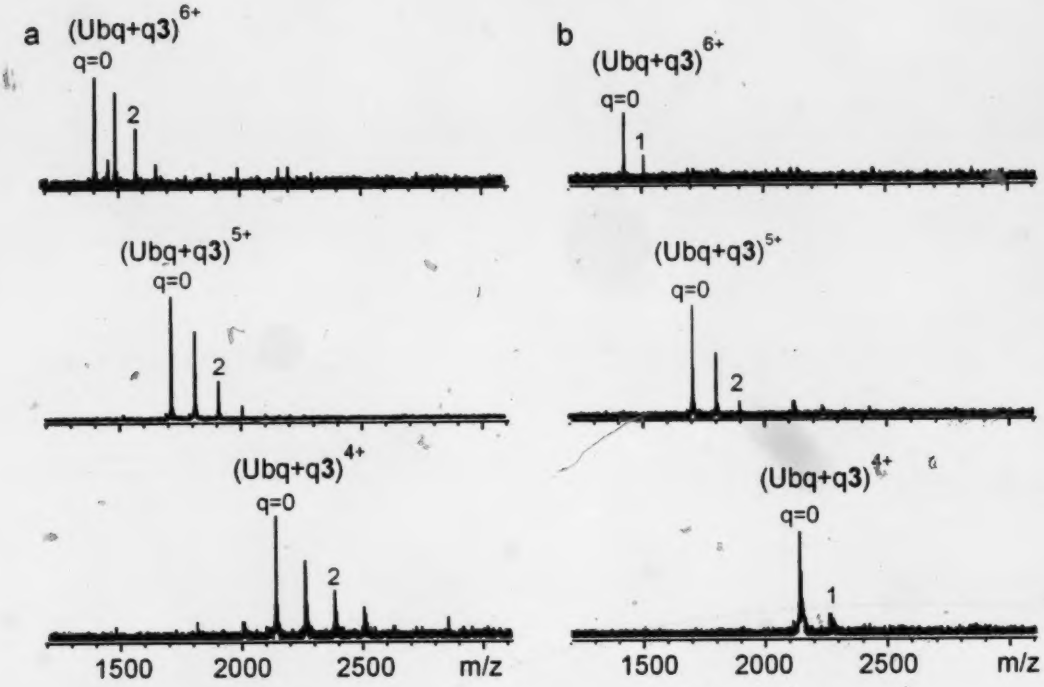


Figure 4.18 (a) Isolation mass spectra obtained from the solutions of 6 μM Ubq and 80 μM **3**; (b) BIRD mass spectra obtained at 160 °C and a reaction time of 1.5 s.

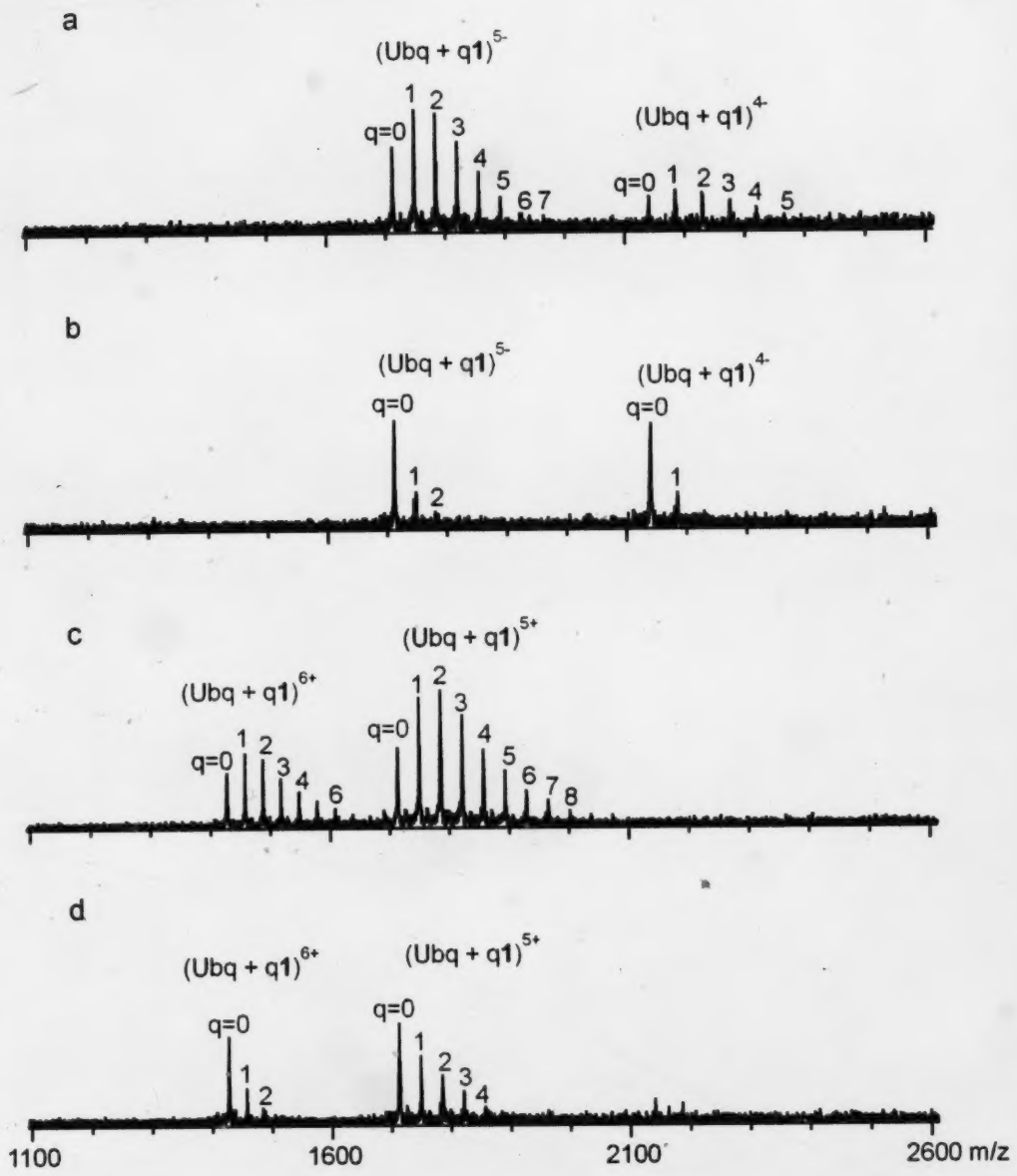


Figure 4.19 NanoES mass spectra obtained from the solutions of 10 μM Ubq and 100 μM **1** in negative ion mode (a) and in positive ion mode (c) and the corresponding BIRD mass spectra obtained at 91 °C and a reaction time of 2 s (b and d).

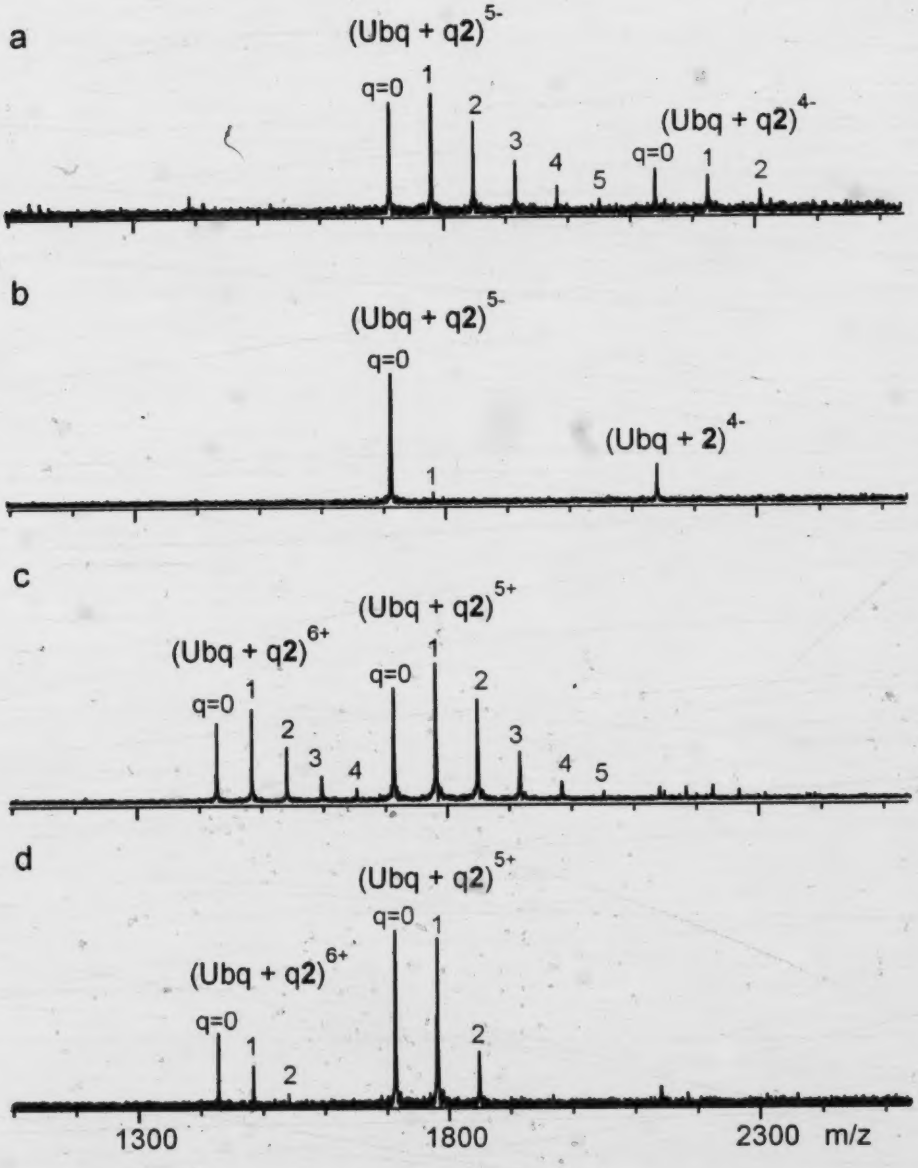


Figure 4.20 NanoES mass spectra obtained from the solutions of 10 μM Ubq and 170 μM **2** in negative ion mode (a) and in positive ion mode (c) and the corresponding BIRD mass spectra obtained at 104 °C and a reaction time of 2 s (b and d).

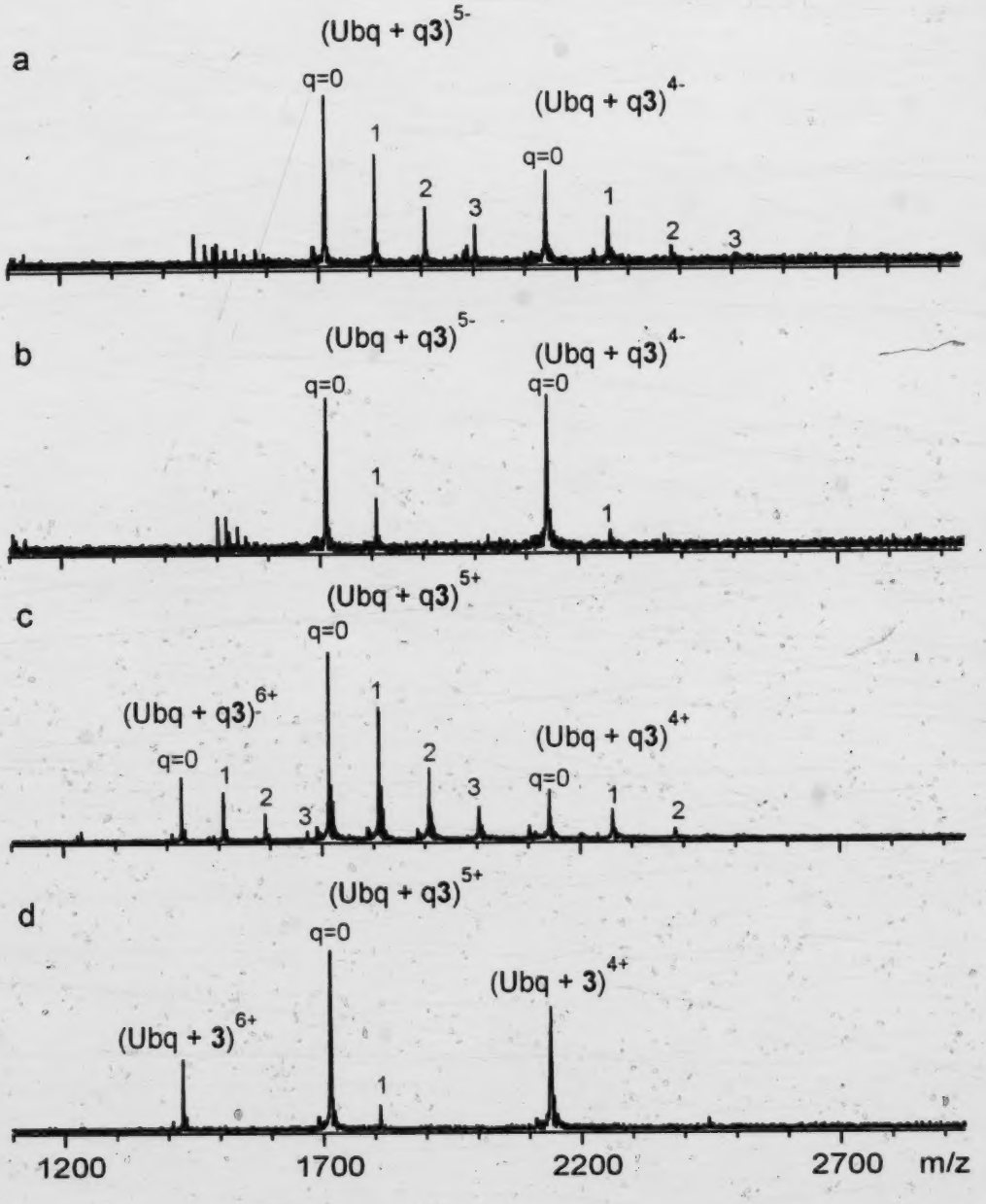


Figure 4.21 NanoES mass spectra obtained from the solutions of 10 μM Ubq and 100 μM **3** in negative ion mode (a) and in positive ion mode (c) and the corresponding BIRD mass spectra obtained at 159 °C and a reaction time of 3 s (b and d).

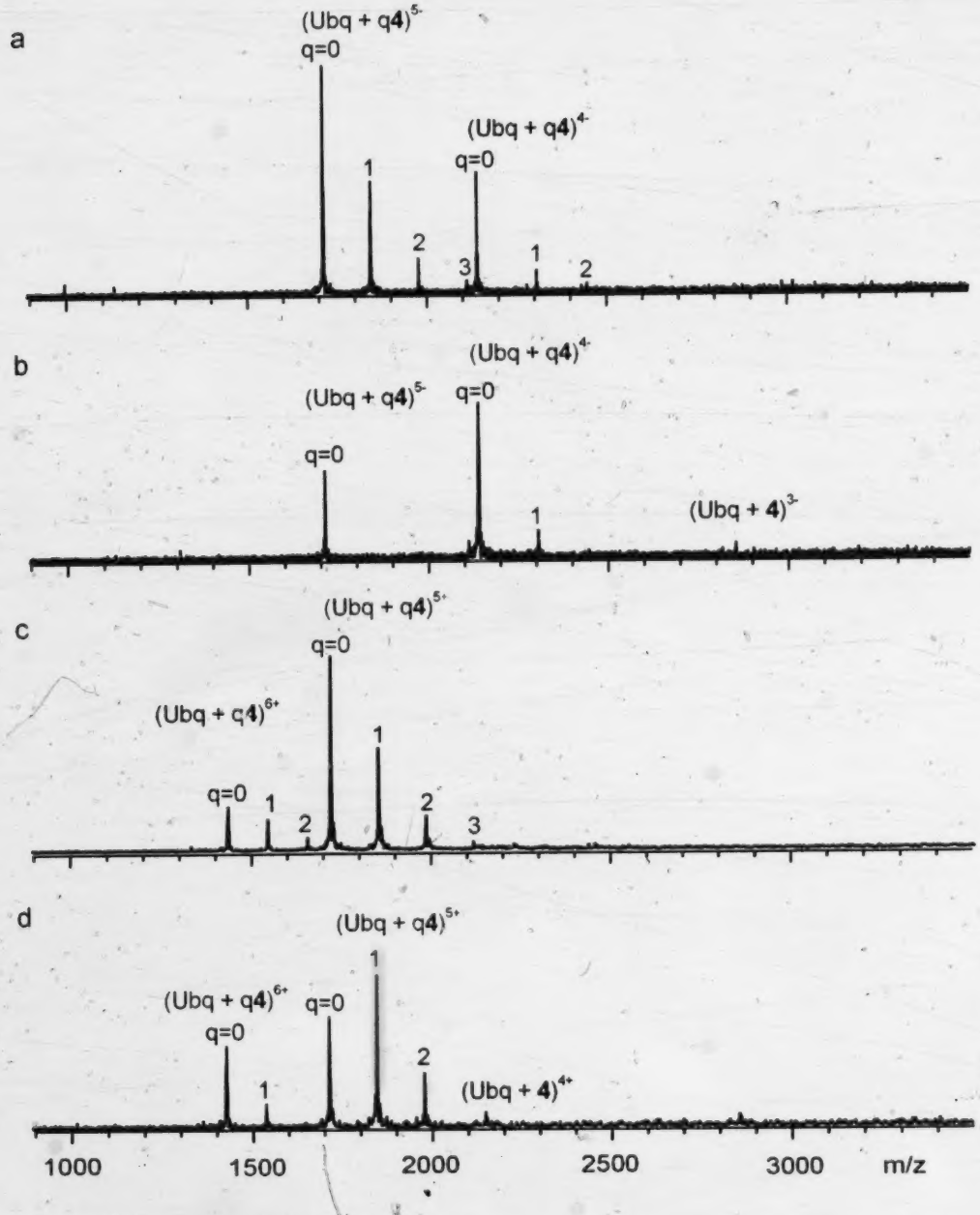


Figure 4.22 NanoES mass spectra obtained from the solutions of 10 μM Ubq and 100 mM **4** in negative ion mode (a) and in positive ion mode (c) and the corresponding BIRD mass spectra obtained at 185 °C and a reaction time of 3 s (b and d).

4.3.2 Minimizing Nonspecific Protein-Carbohydrate Complexes

As stated earlier, one of the objectives of this work was to draw on the insights gained into the factors that influence the formation/stability of the nonspecific complexes to devise new strategies to minimize the appearance of such complexes in the nanoES mass spectra. The observation that the formation of gaseous carbohydrate and carbohydrate cluster ions competes with the formation of nonspecific protein-carbohydrate complexes suggests such a strategy. Enhancement of the processes that lead to the formation of gaseous carbohydrate ions, which can be realized by the addition of metal ions to the nanoES solution should, consequently, reduce the extent of nonspecific protein-carbohydrate binding. To demonstrate the effectiveness of this approach, mass spectra were measured for solutions of CA and **2** in the absence and presence of the metal salts: NaCl, $CaCl_2$ or $MgCl_2$. Shown in Figures 4.23a-d are mass spectra measured for solutions of 6 μM CA, 500 μM **2** and 1 mM CH_3COONH_4 in the absence (Figure 4.23a) and presence of 40 μM NaCl (Figure 4.23b), $CaCl_2$ (Figure 4.23c) or $MgCl_2$ (Figure 4.23d). In each case the solutions containing a metal salt yielded $(CA + q\mathbf{2})^{n+}$ ions with a narrower distribution of q, although the effect was more pronounced for the solutions containing $CaCl_2$ or $MgCl_2$. The values of q_{max} and q_{ave} for the solution without a metal salt were 10 and 4 and these decreased to 8 and 3 with the addition of NaCl, 8 and 2 with $CaCl_2$, and 8 and 2 with $MgCl_2$. Increasing the salt concentration leads to a further reduction in the q_{max} and q_{ave} values. However, this is accompanied by a reduction in the protein and complex ion S/N due, in part, to metal ion adduct formation. The decrease in q_{max} and q_{ave} observed upon addition of metal salts to the solution is attributed to the enhanced conversion of the carbohydrates to gas phase ions, either as monomers or

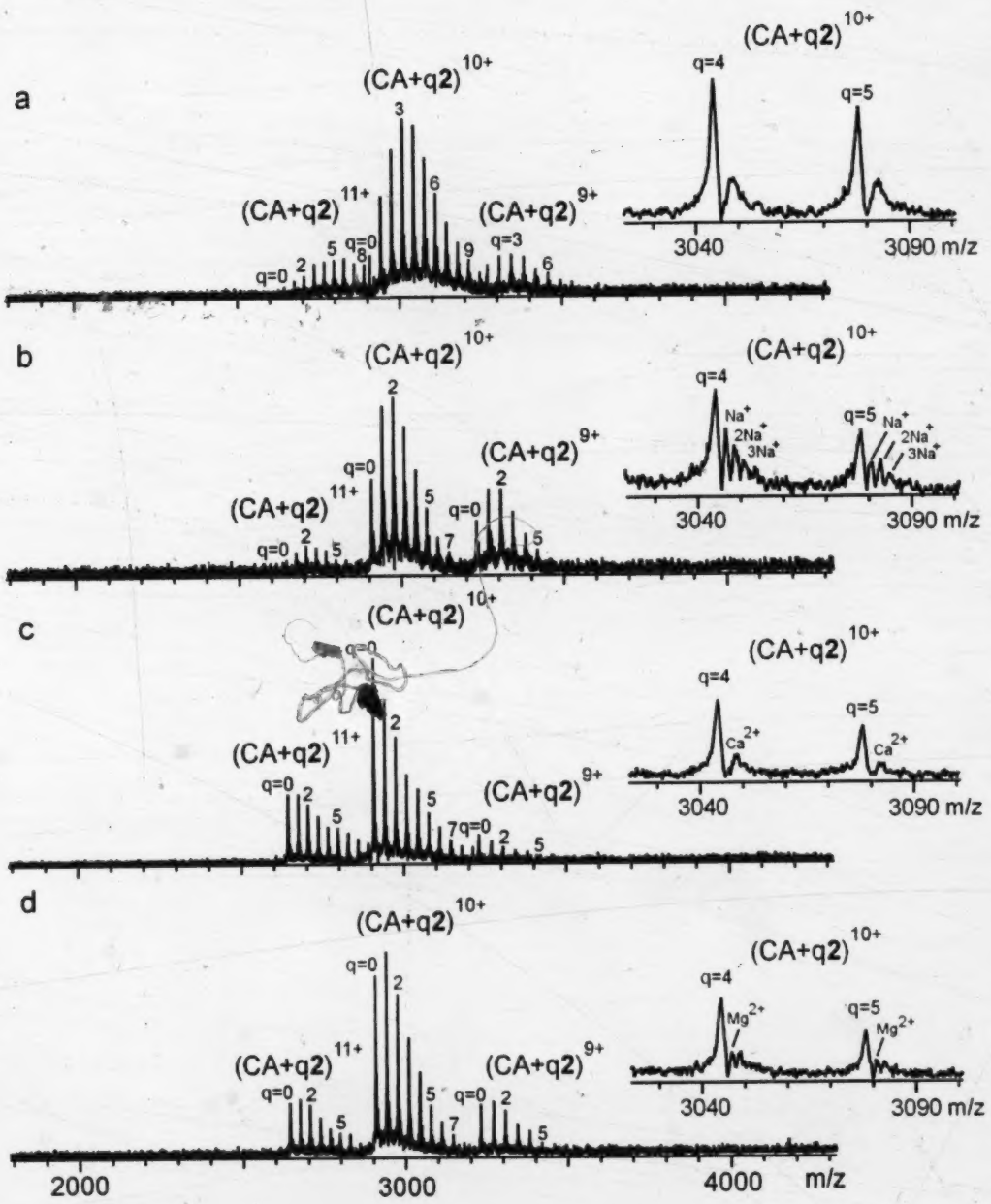


Figure 4.23 NanoES mass spectra obtained from the solutions of 6 μM CA, 500 μM **2** and 1 mM $CH_3COO\ NH_4$ in the absence (a) and presence of 40 μM salt: NaCl (b), $CaCl_2$ (c) or $MgCl_2$ (d).

clusters, by nanoES. The addition of NaCl to the nanoES solution led to more abundant $(L + Na)^+$ ions, while the addition of $CaCl_2$ or $MgCl_2$ led to abundant cluster ions, $(3L + Ca)^{2+}$ and $(3L + Mg)^{2+}$. Although the addition of the metal salts to the nanoES solution does not eliminate the nonspecific complexes, these results clearly demonstrate that this is an effective strategy for reducing nonspecific protein-carbohydrate interactions during the nanoES process.

4.4 Conclusions

Using nanoES-FT-ICR/MS, we have carried out the first comprehensive study of the factors that influence the formation of nonspecific complexes of biopolymers, proteins with carbohydrates, during the nanoES process. Predictably, increasing the concentration of carbohydrate, at a fixed concentration of protein, leads to an increase in the fraction of protein engaged in nonspecific interactions with the carbohydrates. The overall efficiency of nonspecific binding is found to depend on the concentration of the carbohydrate, reaching a maximum value of 5 – 10% over the range of concentrations investigated (23 μM to 900 μM). We propose that the efficiency of the nonspecific binding process and the distribution of bound carbohydrates is governed, predominantly, by the small size of the offspring droplets that ultimately produce the gaseous ions. In addition, the number of carbohydrate ligands bound to protein is influenced by the loss of carbohydrate molecules from the offspring droplets through ion evaporation and the self-association of carbohydrates into clusters within the offspring droplets. Dissociation of weakly interacting protein-carbohydrate complexes in the gas phase prior to detection may also attenuate the degree of nonspecific binding. The structure of the carbohydrate

was found to influence its tendency to engage in nonspecific interactions with the proteins. Small, hydrophilic carbohydrates, which are less prone to ion evaporation from the nanoES droplets and the formation of clusters within the droplets, are more efficient at forming nonspecific complexes compared to larger or hydrophobic carbohydrates. In contrast, the efficiency of nonspecific binding is insensitive to the structure of the protein and to the charge state of the complex. Under certain conditions, the number of bound carbohydrates significantly exceeds the charge state of the complex, indicating that neutral interactions play a significant role in the formation of the nonspecific complexes.

The kinetic stability of the nonspecific complexes was investigated using BIRD. It was found that the stability of the complexes increases with the size of the carbohydrate, a result that can be explained by the greater number of oxygen-containing groups available to participate in intermolecular hydrogen bonds with the protein. The kinetic stability is also influenced by the structure of the protein, with the larger proteins forming more kinetically stable complexes and by the number and the type of charge groups, although no simple relationship between charge state and stability was identified.

Finally, a new strategy, which employs the use of metal salts to minimize the formation of nonspecific protein-carbohydrate complexes during the nanoES process, is proposed. The strategy exploits the competitive nature of gaseous carbohydrate ion formation and the formation of the nonspecific protein-carbohydrate complexes by nanoES. The addition of Na^+, Ca^{2+} and Mg^{2+} to the nanoES solution was shown to enhance the formation of the gaseous carbohydrate ions, either as monomers or clusters, and reduce the degree of nonspecific protein-carbohydrate binding. The divalent Ca^{2+} and

Mg^{2+} ions were found to be more effective than the monovalent Na^{+}, although the effectiveness of all of the salts increased with concentration.

4.5 Literature Cited

(1) (a) Hutchens, T. W.; Allen, M. H.; Li, C. M.;Yip, T. T. *Febs Letters* **1992**, *309*, 170-174; (b) Hu, P. F.; Ye, Q. Z.; Loo, J. A. *Anal. Chem.* **1994**, *66*, 4190-4194; (c) Feng, R.; Castelhano, A. L.; Billedeau, R.; Yuan, Z. Y. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1105-1111; (d) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154-1164; (e) Cheng, X. H.; Harms, A. C.; Goudreau, P. N.; Terwilliger, T. C.; Smith, R. D. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 7022-7027; (f) Loo J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23; (g) Nemirovskiy, O. V.; Ramanathan, R.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 809-812.

(2) (a) Cheng, X.; Chen, R.; Bruce, J. E.; Schwartz, B. L.; Anderson, G. A.; Hofstadler, S. A.; Gale, D. C.; Smith, R. D.; Gao, J.; Sigal, G. B.; Mammen, M.; Whitesides, G. M. *J. Am. Chem. Soc.* **1995**, *117*, 8859-8860; (b) Blair, S. M.; Kempen, E. C.; Brodbelt J. S. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1049-1059; (c) Loo, J. A.; Holler, T. P.; Foltin, S. K.; McConnell, P.; Banotai, C. A.; Horne, N. M.; Mueller, W. T.; Stevenson, T. I.; Mack, D. P. *Proteins: Struct. Funct. Genet.* Suppl **1998**, *2*, 28-37; (d) Rostom, A. A.; Tame, J. R. H.; Ladbury, J. E.; Robinson. C. V. *J. Mol. Biol.* **2000**, *296*, 269-279; (e) Hill, T. J.; Lafitte, D.; Wallace, J. I.; Cooper, H. J.; Tsvetkov, P. O.; Derrick, P. J. *Biochemistry* **2000**, *39*, 7284-7290.

(3) (a) Greig, M. J.; Gaus, H; Cummins, L. L; Sasmor, H.; Griffey, R. H. *J. Am. Chem. Soc.* **1995**, *117*, 10765-10766; (b) Loo, J. A.; Hu, P.; McConnell, P.; Mueller, W. T.; Sawyer, T. K.; Thanabal, V. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 234-243; (c) Jørgensen, T. J. D.; Roepstorff, P.; Heck, A. J. R. *Anal. Chem.* **1998**, *70*, 4427-4432; (d) Ayed, A; Krutchinsky, A. N.; Ens, W.; Standing, K. G.; Duckworth, H. W. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 339-344; (e) Sannes-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264-271; (f) Kempen, E. C.; Brodbelt, J. S. *Anal. Chem.* **2000**, *72*, 5411-5416; (g) Kitova, E. N.; Kitova, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605-611; (h) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945-4955; (i) Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; De Pauw, E. *J. Mass Spectrom.* **2003**, *38*, 491-501; (g) Zhang, S.; Van Pelt, C. K.; Wilson, D. B. *Anal. Chem.* **2003**, *75*, 3010-3018; (k) Hagan, N.; Fabris, D. *Biochemistry*, **2003**, 42, 10736-10745; (l) Tjernberg, A.; Carnö, S.; Oliv, F.; Benkestock, K.; Edlund, P.-O.; Griffiths, W. J.; Hallén, D. *Anal. Chem.* **2004**, *76*, 4325-4331.

(4) Daneshfar, R; Kitova, E. N.; Klassen, J. S. *J. Am. Chem. Soc.* **2004**, *126*, 4786-4787.

(5) Benesch, J. L. P.; Sobott, F.; Robinson, C. V. *Anal. Chem.* **2003**, *75*, 2208-2214.

(6) (a) McCammon, M. G.; Scott, D. J.; Keetch, C. A.; Greene, L. H.; Purkey, H. E.; Petrassi, H. M.; Kelly, J. W.; Robinson, C. V. *Structure* **2002**, *10*, 851-863; (b) Aquilina, J. A.; Benesch, J. L. P.; Bateman, O. A.; Slingsby, C.; Robinson, C. V. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10611-10616; (c) Sobott, F.; Robinson, C. V. *Int. J. Mass Spectrom.* **2004**, *236*, 25-32.

(7) (a) Light-Wahl, K. J.; Schwartz, B. L.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 5271-5278; (b) Nettleton, E. J.; Sunde, M.; Lai, Z.; Kelly, J. W.; Dobson, C. M.; Robinson, C. V. *J. Mol. Biol.* **1998**, *281*, 553-564; (c) Rostom, A. A.; Robinson, C. V. *J. Am. Chem. Soc.* **1999**, *121*, 4718-4719; (d) McCammon, M. G.; Hernández, H.; Sobott, F.; Robinson, C. V. *J. Am. Chem. Soc.* **2004**, *126*, 5950-5951.

(8) (a) Smith D. L.; Zhang Z. *Mass Spectrom. Rev.* **1994**, *13*, 411-429; (b) Smith R. D.; Bruce J. E.; Wu Q.; Lei Q. P. *Chem. Soc. Rev.* **1997**, *26*, 191-202; (c) Loo J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186; (d) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390; (e) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1-27; (d) Heck, A. J. R.; van den Heuvel, R. H. H. *Mass Spectrom. Rev.* **2004**, *23*, 368-389.

(9) (a) Speir, J. P.; Senko, M. W.; Little. D. P.; Loo, J. A.; McLafferty, F. W. *J. Mass Spectrom.* **1995**, *30*, 39-42; (b) Little, D. P.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 209-210; (c) Tolić, L. P.; Bruce, J. E.; Lei, Q. P.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **1998**, *70*, 405-408.

(10) Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. *J. Am. Chem. Soc.* **1996**, *118*, 8646-8653.

(11) Smith, R. D.; Light-Wahl, K. J. *Biol. Mass Spectrom.* **1993**, *22*, 493-501.

(12) (a) Gao, J.; Wu, Q.; Carbeck, J.; Lei, P.; Smith, R. D.; Whitesides, G. M. *Biophysical J.* **1999**, *76*, 3253-3260; (b) Gabelica, V.; Galic, N.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 946-953; (c) Jørgensen, T. J. D.; Hvelplund, P.; Anderson, J. U.; Roepstorff, P. *Int. J. Mass Spectrom.* **2002**, *219*, 659-670.

(13) (a) Zehl, M.; Allmaier, G. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 339-344; (b) van der Kerk-van Hoof, A.; Heck, A. J. R. *J. Mass Spectrom.* **1999**, *34*, 813-819.

(14) Wang, W.; Kitova, E. N.; Klassen, J. S. *J. Am. Chem. Soc.* **2003**, *125*, 13630-13631.

(15) (a) Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644; (b) Dunbar, R. C.; McMahon, T. B. *Science*, **1998**, *279*, 194-197.

(16) Felitsyn, N; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2001**, *73*, 4647-4661.

(17) Wang, W.; Sun, J.; Kitova, E. N.; Klassen, J. S. manuscript is in preparation.

(18) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240-2249.

(19) Peschke, M.; Verkerk, U.; Kebarle, P. *J. Am. Soc. Mass Spectrom.* in press.

(20) Gomez, A.; Tang, K. *Physics of Fluids* **1994**, *6*, 404-414.

(21) (a) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, 64, 2287-2294; (b) Thomson, B. A.; Iribarne, J. V. *J. Chem. Phys.* **1979**, 71, 4451-4463.

(22) (a) Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, *65*, 927A-986A; (b) Tang, L.; Kebarle, P. *Anal. Chem.* **1993**, *65*, 3654-3668.

(23) Schmidt, A.; Karas, M.; Dülcks, T. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 8646-8653.

(24) (a) Meng, C. K.; Fenn, J. B. *Org. Mass Spectrom.* **1991**, *26*, 542-549; (b) Julian, R. R.; Hodyss, R.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2001**, *123*, 3577-3583.

(25) (a) Jurchen; J. C.; Cooper, R. E.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 1477-1487; (b) Jurchen, J. C.; Garcia, D. E.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 1373-1386.

(26) (a) Anacleto, J. F.; Pleasance, S.; Boyd, R. K. *Org. Mass Spectrom.* **1992**, *27*, 660-666; (b) Wang, G.; Cole, R. B. *Anal. Chem.* **1994**, *66*, 3702-3708; (c) Hop, C. E. C. A.; Saulys, D. A.; Gaines, D. F. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 860-865; (d) Zhou, S. L.; Hamburger, M. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 797-800; (e) Wang, G.; Cole, R. B. *Anal. Chem.* **1998**, *70*, 873-881; (f) Gamero-Castano, M.; de la Mora, J. F. *J. Mass Spectrom.* **2000**, *35*, 790-803; (g) Kebarle, P.; Peschke, M. *Anal. Chim. Acta* **2000**, *406*, 11-35; (h) Gamero-Castano, M.; de la Mora, J. F. *Anal. Chim. Acta* **2000**, *406*, 67-91.

(27) Hofstadler, S. A.; Sannes-Lowery, K. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1-9.

(28) (a) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531-544; (b) Di Stefano, D. L.; Wand, A. J. *Biochemistry* **1987**, *26*, 7272-7281; (c) Pan, Y.; Briggs, M. S. *Biochemistry* **1992**, *31*, 11405-11412.

(29) The molecular surface area calculations were performed using Hyperchem version 7.51 (Hypercube Inc, Gainsville, FL). A solvent probe radius of 1.4 Å was used. The crystal structure of the human CA II isozyme was used to estimate the surface area of bovine CA II.

(30) (a) Clemmer, D. E.; Hudgins, R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1995**, *117*, 10141-10142; (b) Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 2240-2248.

Chapter 5

Nonspecific Protein-Carbohydrate Complexes Produced by Nanoelectrospray Ionization. Investigating the Nature of the Noncovalent Interactions

5.1 Introduction

The application of MS to biochemical problems has been revolutionized by the development of the ES ionization technique, wherein dissolved species, which may be neutral or carry a net charge, are transferred into the gas phase in an ionized form. An important feature of the ES technique is the ability to transfer intact noncovalently-bound biological complexes, such as protein-receptor, enzyme-substrate and multiprotein complexes, from buffered aqueous solutions to the gas phase. As a result, ES-MS has become a powerful tool for detecting specific biomolecular complexes in solution and is increasingly used to quantify their binding stoichiometry and affinity [1-6]. While solution specific biomolecular complexes are often easily detected by ES-MS, it is not clear to what extent the structures of the gaseous complexes and, in particular, the intermolecular interactions resemble those present in solution. Nonspecific interactions, *i.e.* interactions not present in solution and which form during or after the ES processes, have been shown to contribute significantly to the stability of some specific complexes in the gas phase [7-11]. The formation of nonspecific interactions during the ES process may be beneficial and, perhaps, even necessary to the survival of certain biological complexes, in particular those stabilized predominantly by hydrophobic interactions [7]. In addition to the formation of nonspecific interactions within specific complexes,

nonspecific interactions between biomolecules that do not bind in solution and between biomolecules and small neutral molecules or ions present in solution may also occur during the ES process. Although the mechanism of their formation is not fully understood, nonspecific complexes involving macromolecules, such as proteins, are widely believed to form according to the charge residue ES model (CRM), whereby the nonspecific intermolecular interactions occur as the charged ES nanodroplets (offspring droplets), containing multiple analytes, evaporate to dryness [12]. The gaseous complexes produced by the nonspecific interaction may be sufficiently stable (kinetically) that they survive the ion source and are detected.

Although the tendency of biological molecules to engage in nonspecific intermolecular interactions during the ES process, particularly at elevated concentrations, is widely recognized there is little known about the nature of the interactions responsible for stabilizing the gaseous complexes. A greater understanding of the nonspecific complexes, their formation, structure and stability, may facilitate the development of new strategies to reduce or minimize the appearance of these complexes in ES spectra. Identifying and quantifying the intermolecular interactions responsible for stabilizing nonspecific complexes may also reveal new insights into the structural changes that accompany the desolvation of biological complexes during ES and, thereby, aid the development of direct ES-MS-based techniques to characterize the structures of biomolecular complexes in solution. Gas phase studies of nonspecific interactions between proteins and small polyfunctional molecules, such as polyols and amino acids, are also relevant to ongoing efforts to elucidate the mechanisms by which these

molecules are able to stabilize proteins and biologicals in low humidity environments [13].

To develop a greater understanding of factors responsible for the formation of nonspecific complexes between proteins and polyfunctional molecules, including biopolymers, and the forces that stabilize the gaseous complexes, our laboratory has undertaken a comprehensive investigation of nonspecific protein-carbohydrate complexes originating from nonspecific interactions formed by nanoES. In the first of a series of studies demonstrated in Chapter 3, the kinetic and energetic stability of two gaseous complexes composed of a 26.5 kDa single chain variable fragment (scFv) of a monoclonal antibody and the αGal[αAbe]αMan trisaccharide were investigated using the BIRD technique, implemented with FT-ICR/MS. One of the gaseous complexes was produced by transferring the specific complex from solution to the gas phase with nanoES. The second complex was produced by nonspecific interactions during the nanoES process. This study revealed that, at the charge states investigated (+10 and +11), the nonspecific complex was kinetically more stable and, in the case of the +10 charge state, energetically more stable than the specific complex. We proposed two explanations for the enhanced stability of the nonspecific complex, relative to the specific complex. We suggested that conformational constraints imposed by the specific interactions in solution prevented the specific complex from adopting a lower energy structure in the gas phase, while the nonspecific complex, unencumbered by such constraints, is able to explore a greater region of conformational space and relax to a lower energy structure. As an alternative explanation, we suggested that strong ionic hydrogen bonding, which is believed to be absent in the specific complex, might contribute to the stability of the

nonspecific complex. The ionic hydrogen bond(s) in this case would likely involve the solvation of a protonated basic residue, such as arginine or histidine, by one or more carbohydrate hydroxyl groups. Support for the latter explanation was found in a comparison of Arrhenius parameters determined for the nonspecific scFv•αGal [αAbe]αMan and a second nonspecific complex, CA•αGal[αAbe]αMan (CA = bovine carbonic anhydrase II) at the same charge states. The Arrhenius parameters for both complexes were indistinguishable, within experimental error, suggesting similar intermolecular interactions in both complexes. The absence of a dependence of carbohydrate binding on protein structure (primary and higher order) is consistent with the contribution of strong ionic hydrogen bonds to the stability of both complexes.

In Chapter 4, we employed nanoES-FT-ICR/MS and BIRD to investigate the nonspecific interactions between CA and bovine ubiquitin (Ubq) with a series of carbohydrates, ranging in size from mono- to tetrasaccharides. The goal of this study was to identify the factors that influence the tendency of proteins and carbohydrates to engage in long-lived nonspecific interactions during the nanoES process. Notably, it was found that the degree of nonspecific binding (*i.e.* the number of carbohydrates bound to a given protein) is insensitive to the charge state of the complex and, at high carbohydrate concentrations, the number of carbohydrates bound to the protein can significantly exceed the charge state. Based on these results it was concluded that formation of the nonspecific complexes is not governed simply by charge solvation and that neutral interactions also contribute significantly. However, BIRD experiments performed simultaneously on the protein-carbohydrate complexes with a distribution of bound carbohydrates revealed that, in addition to being sensitive to the structure of the protein

and carbohydrate, the number and type of charged groups influences the kinetic stability of the complexes. Interestingly, no simple relationship between charge state and stability was evident.

The aforementioned studies provide indirect evidence that both neutral and ionic intermolecular interactions play a role in stabilizing the nonspecific protein-carbohydrate complexes in the gas phase. However, the contribution of these interactions, and the influence of the charge state, the structure of the protein and carbohydrate, and the number of bound carbohydrates thereon, remains unclear. In an effort to more fully evaluate the nature of the intermolecular interactions responsible for the nonspecific complexes, we have carried out the first comprehensive study of the dissociation kinetics and energetics of protein-ligand complexes arising from nonspecific interactions during nanoES. Using BIRD, time-resolved thermal dissociation experiment were performed on a series of 1:1 protein-carbohydrate complexes formed between the proteins CA, Ubq and bovine pancreatic trypsin inhibitor (BPTI) and carbohydrates ranging in size from mono- to tetrasaccharide. From an analysis of the Arrhenius parameters determined for a series of protonated 1:1 protein-trisaccharide complexes, the influence of charge state, protein structure and carbohydrate structure on the kinetic and energetic stability of the nonspecific complexes was assessed. BIRD was also performed on a number of 1:2 protein-trisaccharide complexes to establish whether the proteins present a unique and independent carbohydrate binding site in the gas phase.

5.2 Experimental

5.2.1 Proteins and Carbohydrates

Bovine carbonic anhydrase II, CA (molecular weight ≡ MW 29 088 Da), bovine ubiquitin, Ubq (MW 8 565 Da), and bovine pancrease trypsin inhibitor, BPTI (MW 6 517 Da), were purchased from Sigma Canada and used without further purification. The monosaccharide D-Gal (**1**) was purchased from Sigma Canada and the carbohydrates, αAbe(2-O-CH_3-αMan) (**2**), αTal[αAbe]αMan (**3**), αAbe(2-O-CH_3-αMan) αGlcαGlc (**4**), αGal[αAbe]αMan (**5**), αGal[αAbe](4-deoxyαMan) (**6**) and (6-deoxyαGal)[αAbe]αMan (**7**) were provided by D. R. Bundle (Univ. of Alberta). The structures of the carbohydrates are shown in Figure 5.1.

5.2.2 Mass Spectrometry

Gaseous complexes were produced by nanoES performed on aqueous solutions of protein (~10 μM), carbohydrate (40 - 80 μM) and ammonium acetate (1 mM). The nanoES tips were constructed from aluminosilicate capillaries (1mm o.d., 0.68 mm i.d.), pulled to approximately 4 - 7 μm o.d. at one end using a P-2000 micropipette puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution was established by applying a voltage of ± 800-1000 V to a platinum wire inserted inside the glass tip. The solution flow rate was typically 20 to 50 nL/min. The droplets and gaseous ions produced by nanoES were introduced into the vacuum chamber of a modified ApexII 47e Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker, Billerica MA) through a stainless steel capillary (i.d. 0.43 mm) maintained at an external temperature of 66 °C. The ion/gas jet sampled by the capillary (± 48 - 52 V) was transmitted through a skimmer

(± 0 - 2 V) and stored, in a hexapole. Ions were accumulated in the hexapole for 2s, then ejected and accelerated (± 2700 V) into a 4.7 tesla superconducting magnet, decelerated and introduced into the ion cell. The trapping plates of the cell were maintained at a constant potential of ± 1.4 - 1.8 V throughout the experiment. The typical base pressure for the instrument was ~$5x10^{-10}$ mbar. The temperature of the ion cell was controlled with two external flexible heating blankets placed around the vacuum tube in the vicinity of the cell [14].

Data acquisition was controlled by an SGI R5000 computer running the Bruker Daltonics XMASS software, version 5.0. Mass spectra were obtained using standard experimental sequences with chirp broadband excitation. Isolation of the reactant ions for the BIRD experiments was achieved using a combination of single rf frequency and broadband rf sweep excitation. The isolated ions were stored inside the heated cell for varying reaction times prior to excitation and detection. The time-domain spectra, consisting of the sum of 30-60 transients containing 128 K data points per transient, were subjected to one zero-fill prior to Fourier-transformation.

5.3 Results and Discussion

5.3.1 Formation of 1:1 and 1:2 Nonspecific Protein-Carbohydrate Complexes

As demonstrated in Chapter 2, nonspecific protein-carbohydrate complexes are readily produced by nanoES performed on solutions of protein and carbohydrate, with an initial concentration ratio of [carbohydrate]/[protein] > 2. In the present work, aqueous solutions containing 10 μM protein (CA, Ubq or BPTI) and 40 - 80 μM carbohydrate (**1** – **7**) were used to generate protonated and deprotonated nonspecific 1:1 and 1:2 protein-

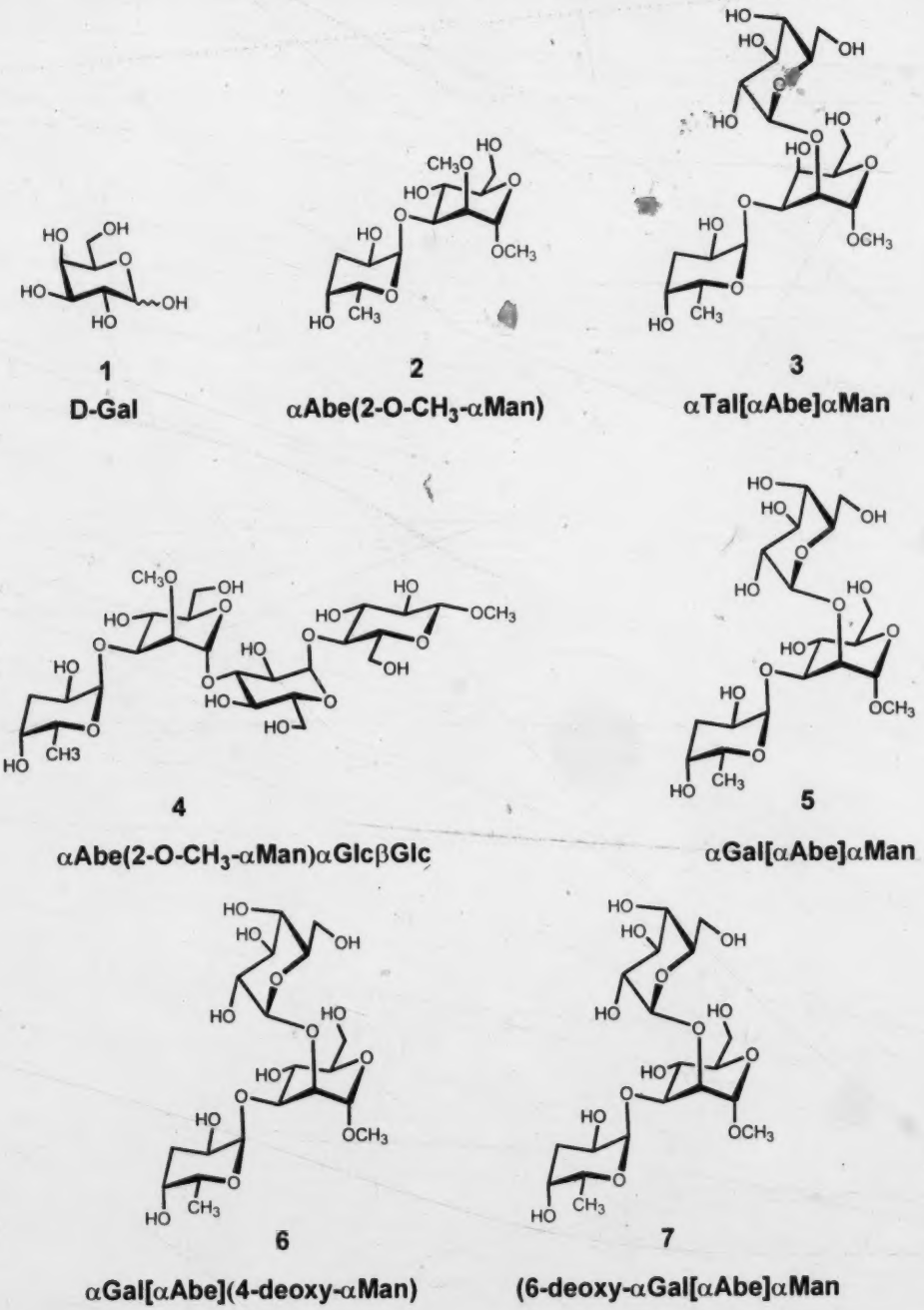


Figure 5.1 Structures of the carbohydrates: D-Gal (**1**), αAbe(2-O-CH_3-αMan) (**2**), αTal[αAbe]αMan (**3**), αAbe(2-O-CH_3-αMan)αGlcβGlc (**4**), αGal[αAbe]αMan (**5**), αGal[αAbe](4-deoxyαMan) (**6**), (6-deoxyαGal)[αAbe]αMan (**7**),

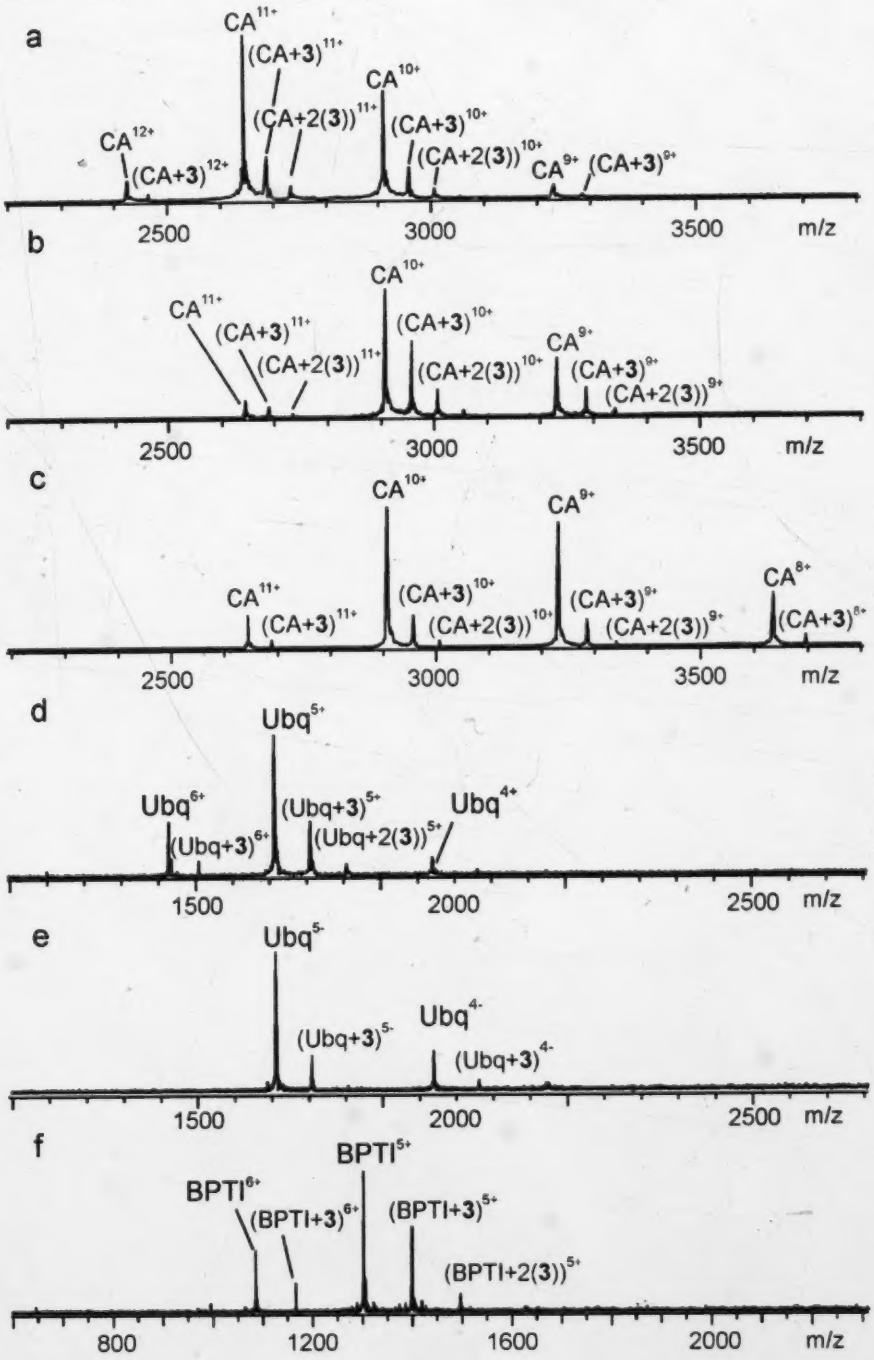

a
CA11+
(CA+3)11+
CA10+
(CA+2(3))11+
(CA+3)10+
CA12+
(CA+3)12+
(CA+2(3))10+
CA9+
(CA+3)9+
2500
3000
3500
m/z
b
CA10+
CA11+
(CA+3)10+
(CA+3)11+
CA9+
(CA+2(3))11+
(CA+2(3))10+
(CA+3)9+
(CA+2(3))9+
2500
3000
3500
m/z
c
CA10+
CA9+
CA11+
(CA+3)10+
(CA+3)9+
CA8+
(CA+3)11+
(CA+2(3))10+
(CA+2(3))9+
(CA+3)8+
2500
3000
3500
m/z
d
Ubq5+
Ubq6+
(Ubq+3)5+
Ubq4+
(Ubq+3)6+
(Ubq+2(3))5+
1500
2000
2500
m/z
e
Ubq5-
Ubq4-
(Ubq+3)5-
(Ubq+3)4-
1500
2000
2500
m/z
f
BPTI5+
BPTI6+
(BPTI+3)6+
(BPTI+3)5+
(BPTI+2(3))5+
800
1200
1600
2000
m/z

Figure 5.2 NanoES mass spectra obtained in the positive ion mode from solutions of protein and **3**, (a) 10 μM CA and 47 μM **3**; (b) 10 μM CA and 71 μM **3**; (c) 10 μM CA and 47 μM **3** with 0.1 mM imidazole; (d) 10 μM Ubq and 45 μM **3**; (f) 10 μM BPTI and 63 μM **3**; (e) nanoES mass spectrum obtained in the negative ion mode from a solution of 10 μM Ubq and 41 μM **3**.

carbohydrate complexes by nanoES. Shown in Figure 5.2 are nanoES mass spectra obtained from the solutions of **3** with CA and BPTI in positive ion mode and with Ubq in positive and negative ion mode. At these concentrations, the major ions observed in positive ion mode correspond to protonated unbound protein, $(P + nH)^{n+} \equiv P^{n+}$, the 1:1 protein-carbohydrate complex, $(P + L + nH)^{n+} \equiv (P + L)^{n+}$, and the 1:2 protein-carbohydrate complex, $(P + 2L + nH)^{n+} \equiv (P + 2L)^{n+}$. CA ions are typical observed at charge states of +9 to +12, while the Ubq and BPTI ions are observed at +5 and +6. A modest shift in the charge state distribution of the protein and complex ions could be achieved through the addition of imidazole to the nanoES solution. Imidazole is a relatively strong base in the gas phase (gas phase basicity = 217 kcal/mol [15]) and can effect proton abstraction from the protonated protein and complex ions in the source region. For example, the addition of 0.1 mM imidazole to solution of CA and carbohydrate shifts the protein/complex ions charge state distribution to +8 - +10, see Figure 5.2c. NanoES mass spectra were also measured for a solution of Ubq and **3** in negative ion mode. The major species observed are the deprotonated protein and 1:1 complex, P^{z-} and $(P+L)^{z-}$, at charge states z = 4 – 5 (Figure 5.2e).

5.3.2 Dissociation Pathways

BIRD experiments were performed on the protonated 1:1 complexes of CA with **1**, **2**, **4** - **7** at charge states n = 10, 11 and with **3** at charge states n = 8 - 12, the 1:1 complexes of Ubq with **1** - **5** at n = 5 and with **3** at z = 5 and the 1:1 complex of BPTI with **5** at n = 5. BIRD was also performed on the protonated 1:2 complexes of **3** with CA at n = 10, 11 and Ubq at n = 5. Over the range of temperatures investigated, 60 –190 °C, BIRD of the $(P + L)^{n+/z-}$ ions proceeds exclusively by the dissociation of the noncovalent

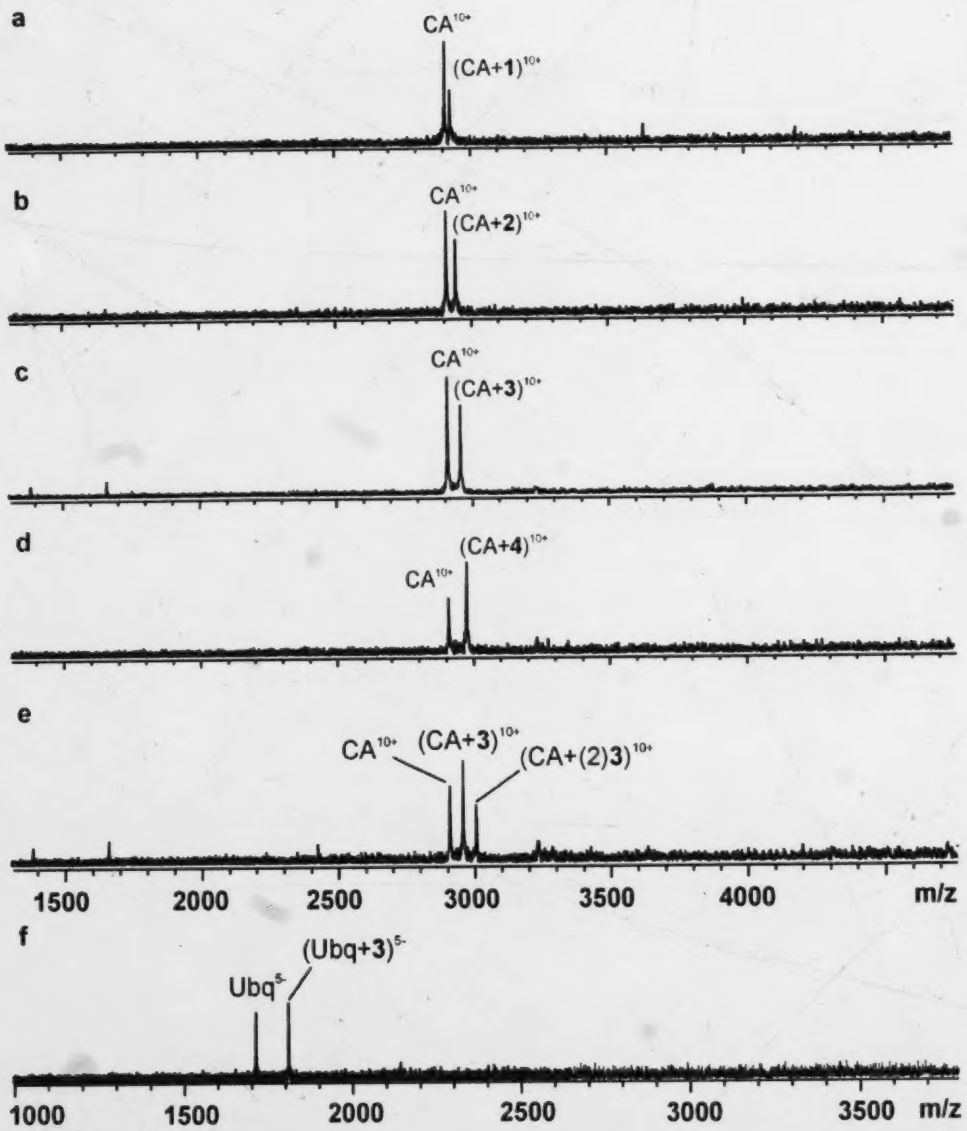


Figure 5.3 BIRD mass spectra obtained for: (a) $(CA + \mathbf{1})^{10+}$, at a temperature of 100 °C and a reaction time 1.5s; (b) $(CA + \mathbf{2})^{10+}$, 128 °C and 1.5s; (c) $(CA + \mathbf{3})^{10+}$, 171 °C and 1.5s; (d) $(CA + \mathbf{4})^{10+}$, 187 °C and 6s; (e) $(CA + (2)\mathbf{3})^{10+}$,171 °C and 1.5s; (f) $(Ubq + \mathbf{3})^{5-}$,163 °C and 1.5s.

intermolecular interactions, with the carbohydrate lost as a neutral, eq. 5.1. From double resonance experiments, wherein suspected reaction intermediates are continuously ejected from the ion cell during BIRD, it was conclusively established that dissociation of the protonated 1:2 complexes, $(P + 2L)^{n+}$, proceeds exclusively by the sequential loss of neutral carbohydrate molecules (eq. 5.2). Illustrative BIRD mass spectra, obtained for the $(CA + L)^{10+}$ ions, where L = **1** – **4** and the $(Ubq + \mathbf{3})^{5+}$ and $(CA + (2)\mathbf{3})^{10+}$ ions, are shown in Figure 5.3

$$(P + L)^{n+/z-} \rightarrow P^{n+/z-} + L \tag{5.1}$$

$$(P + 2L)^{n+} \rightarrow (P + L)^{n+} + L \rightarrow P^{n+} + 2L \tag{5.2}$$

5.3.3 Dissociation Kinetics

The kinetic data for the dissociation of the 1:1 and 1:2 complexes were analyzed by plotting the natural logarithm of the normalized abundance of the complex, A_{PLq}, versus reaction time (t). A_{PLq} was calculated using eq. 5.3:

$$A_{PLq} = I_{PLq} / \sum_{q} I_{PLq} \tag{5.3}$$

where I_{PLq} is the measured ion intensity of the protein bound to q carbohydrates. Over the range of temperatures investigated, dissociation of the 1:1 complexes containing a tri- or tetrasaccharide (**3** – **7**) proceeds by simple first order kinetics, *i.e.* plots of $\ln A_{PLq}$ versus t are linear and the slope is equal to the negative of the rate constant, k:

$$\ln(A_{PLq}) = -kt \tag{5.4}$$

Furthermore, the kinetic data are highly reproducible and insensitive to the protein and carbohydrate concentrations or nanoES spray/source conditions used. Illustrative plots of the kinetic data measured for the $(CA + L)^{10+}$ and $(Ubq + L)^{5+}$ ions, where L = **3**, **4**, are shown in Figures 5.4a and 5.5a. In contrast, the rate of dissociation of the

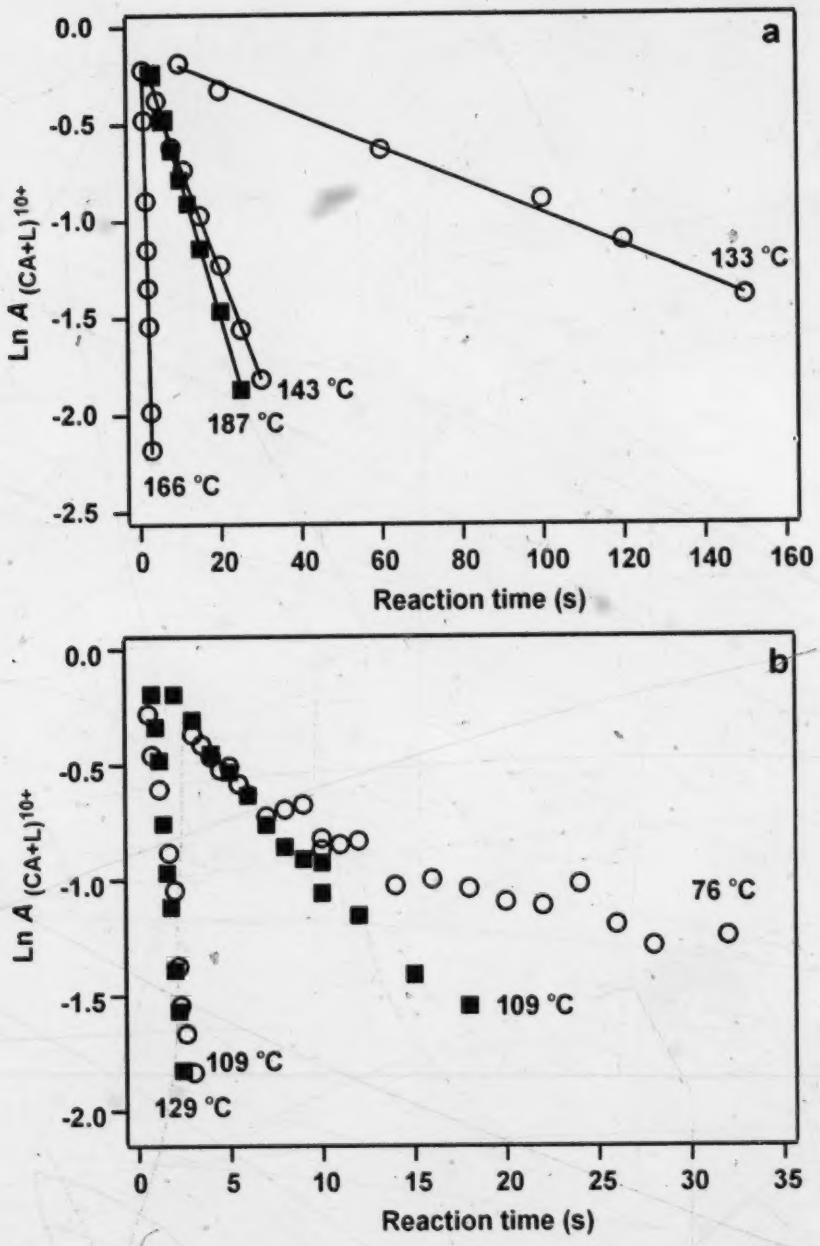


Figure 5.4 Plots of the natural logarithm of the normalized abundance of the complex ion, $(CA + L)^{10+}$, where L = **1** – **4**, versus reaction time at the temperatures indicated: (a) ○, L = **3**; ■, L = **4**; (kinetic data fit to first order kinetics) (b) ○, L = **1**; ■, L = **2**.

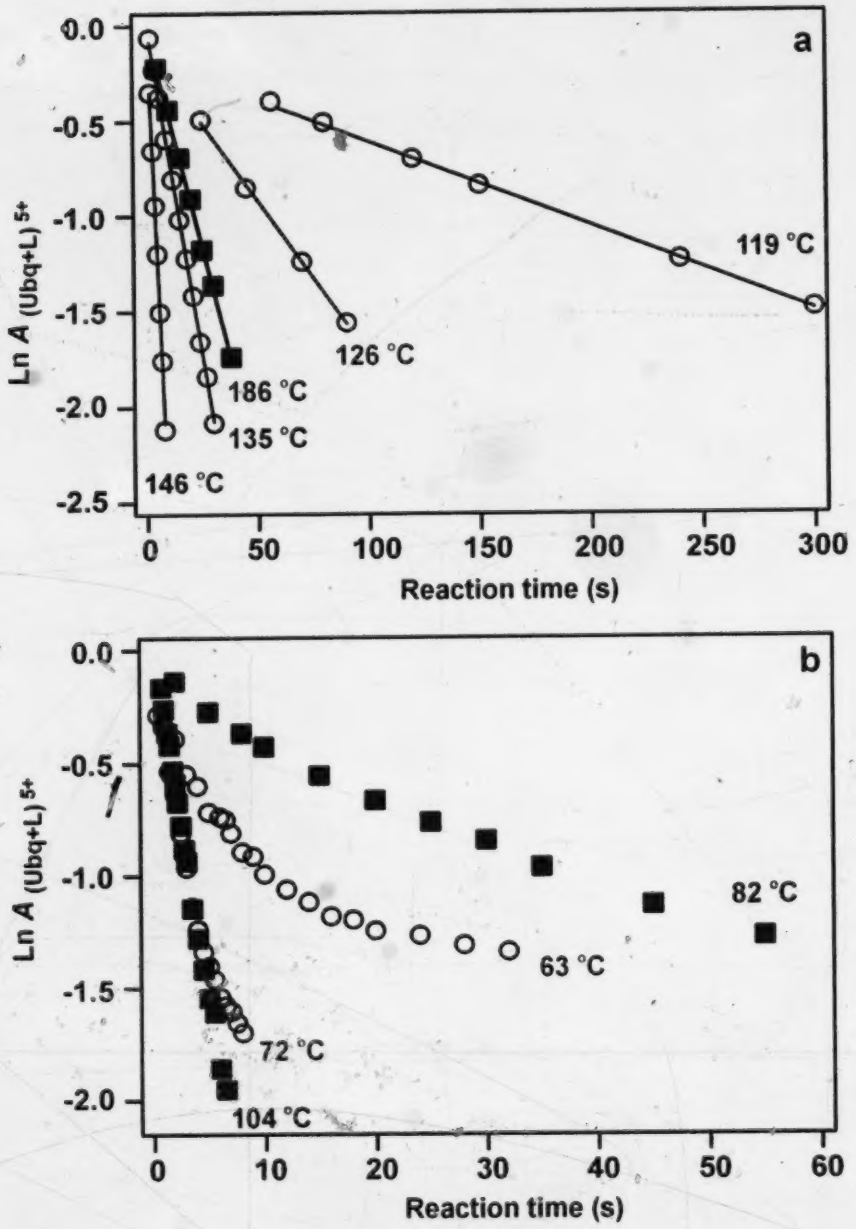


Figure 5.5 Plots of the natural logarithm of the normalized abundance of the complex ion $(Ubq + L)^{5+}$, where L = **1** – **4**, versus reaction time at the temperatures indicated: (a) ○, L = **3**; ■, L = **4**; (kinetic data fit to first order kinetics) (b) ○, L = **1**; ■, L = **2**.

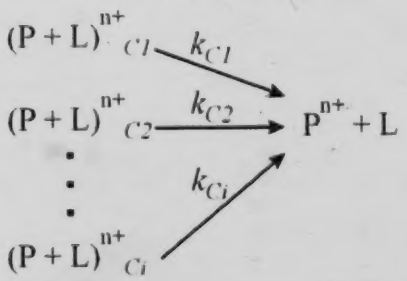
complexes containing the mono- (1) or disaccharide (2) decreases with reaction time; the magnitude of change in reaction rate being most significant at the lowest reaction temperatures studied. In addition, the kinetic data tend to be much less reproducible than is the case for the complexes of **3** - **7**. Kinetic data measured for the $(CA + L)^{10+}$ and $(Ubq + L)^{5+}$ ions, where L = **1**, **2**, are shown in Figures 5.4b and 5.5b.

There are a number of possible explanations for the nonlinear first-order kinetic plots obtained for the complexes of **1** and **2**. If a given $(P + L)^{n+}$ complex consists of multiple, non-interconverting conformers (*i.e.* $(P + L)^{n+}{}_{C1}$, $(P + L)^{n+}{}_{C2}$,...), with distinct dissociation rate constants, k_{C1}, k_{C2}, ... (Scheme 5.1), the observed rate of dissociation will reflect the reactivity of all of the conformers as shown in eq. 5.5:

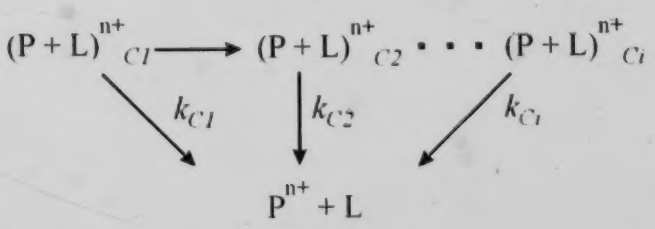
$$A_{PL} = f_{C1}\exp(-k_{C1}t) + f_{C2}\exp(-k_{C2}t) + \bullet\bullet\bullet + f_{Ci}\exp(-k_{Ci}t) \qquad (5.5)$$

where f_C refers to the fractional abundance of a given conformer. Because of differences in the magnitude of the rate constants, the rate of dissociation will decrease with time due to the enhanced decay of the more reactive species. The contribution of kinetically distinct structures is a reasonable explanation given that the nonspecific binding process in ES, by its very nature, is expected to produce complexes with many different carbohydrate interaction sites (at least at the time of their formation). An alternative explanation that can not be discounted is the possible influence of thermally driven changes in the structure of the complex which may occur over the course of reaction and, thereby, alter the rate of dissociation, Scheme 5.2. If such structural changes do indeed occur, the decrease in reaction rate with time indicates that the complexes become kinetically more stable with reaction time, presumably due to the formation of additional or stronger intermolecular interactions, *i.e.* the complex relaxes to a lower energy state.

Finally, changes in reaction rate in BIRD experiments can, in principle, arise from changes in the internal energy distribution of the reactant ions over the course of the reaction. Such a change in the energy distribution is expected when the ions are not in the rapid energy exchange (REX) limit; ions with internal energies greater than the dissociation threshold energy for dissociation will become depleted by dissociation, resulting in a truncated Boltzmann distribution of internal energy [16]. However, this explanation seems unlikely given that BIRD experiments performed on a number of small ions that were not in the REX limit yielded simple first-order kinetics [17], indicating that the internal energy distribution of the ions very rapidly reaches a steady state.



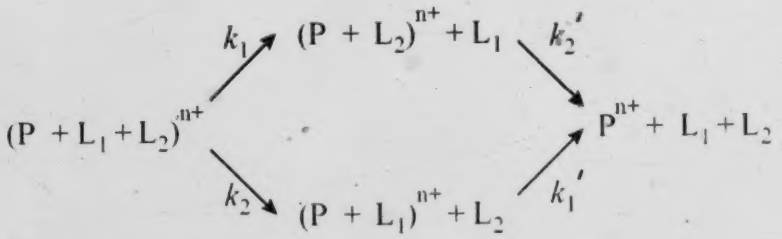
Scheme 5.1 Dissociation pathways of protein-ligand complex ions, $(P + L)^{n+}$, which consists of multiple conformers with distinct dissociation rate constants.



Scheme 5.2 Dissociation pathways of protein-ligand complex, $(P + L)^{n+}$ which undergoes conformational changes over the course of the reaction.

Perhaps more intriguing than the nonlinear kinetics plots obtained for complexes of **1** and **2** are the linear plots obtained for the 1:1 complexes of **3** – **7**. The simple first-order kinetics are suggestive of gaseous $(P + L)^{n+}$ ions with a single structure (in terms of the intermolecular interactions) or multiple kinetically-equivalent structures. As discussed above, the nonspecific association process in ES is expected to generate a number of structurally distinct complexes in which the carbohydrate is bound at different sites. One possible explanation for the observed dissociation kinetics is that the nonspecific complexes, independent of their original structure (*i.e.* carbohydrate binding site), rapidly relax to a single structure or multiple equivalent structures in the gas phase. In other words, the proteins have preferred carbohydrate binding site(s) in the gas phase which the carbohydrate molecules are able to rapidly access (relative to the time scale of the experiment). In an effort to establish whether the nonspecific protein-trisaccharide complexes adopt a single carbohydrate binding site or whether multiple equivalent binding sites exist, BIRD was performed on the nonspecific 1:2 complexes of CA and Ubq with **3** and the results compared with the kinetic data of the corresponding 1:1 complexes. In the simplest binding mode for the 1:2 complex, each carbohydrate occupies a distinct binding site on the protein; these binding sites may or may not be kinetically equivalent and may or may not be independent of one another. Dissociation of the complex will proceed via parallel pathways involving the loss of either of the carbohydrates, L_1 and L_2 (Scheme 5.3) and the measured dissociation rate constant (k_{obs}) will be equal to the sum of the individual dissociation rate constants, *i.e.* $k_{obs} = k_1 + k_2$. If the binding sites are equivalent (kinetically) and independent, k_{obs} will be equal to twice

the rate constant ($k_{1:1}$) measured for the dissociation of the 1:1 complexes, *i.e.* k_{obs} $=2(k_{1:1})$. If the binding sites are nonequivalent but independent, then $k_{obs} \neq 2(k_{1:1})$. Shown



Scheme 5.3 Sequential dissociation pathways of a 1:2 protein-carbohydrate complex $(P + L_1 + L_2)^{n+}$. L_1 and L_2 correspond to the same carbohydrate bound at two different sites.

in Figure 5.6 are plots of $\ln(A_{PLq})$ versus t obtained for the $(CA + 2(\mathbf{3}))^{10/11+}$ and $(Ubq + 2(\mathbf{3}))^{5+}$ ions. It can be seen that the kinetic plots obtained for the three complex ions are linear over the range of temperatures investigated, allowing for the accurate determination of k_{obs}. Plots of the ratio $k_{obs}/k_{1:1}$ versus reaction temperature for the three ions (Figure 5.7) demonstrate that the 1:2 complexes are kinetically less stable than the corresponding 1:1 complexes over the temperature range investigated. For the $(CA + 2(\mathbf{3}))^{10+}$ and $(Ubq + 2(\mathbf{3}))^{5+}$ ions, $k_{obs}/k_{1:1} \geq 2$, while for $(CA + 2(\mathbf{3}))^{11+}$, $k_{obs}/k_{1:1} < 2$. For all three ions, the ratio decreases with increasing temperature. Taken on their own, these results suggest that the carbohydrate binding sites are not equivalent. However, this interpretation assumes that the binding sites are independent of one another. To establish the independence of carbohydrate binding sites, the change in relative abundance of the 1:2 complex ($(P + L_1 + L_2)^{n+}$), the 1:1 complex (the sum of the intermediates $(P + L_1)^{n+}$ and $(P + L_2)^{n+}$) and the 1:0 species (the unbound protein, (P^{n+})), with t was compared to values calculated according to different kinetic models. Three different kinetic models,

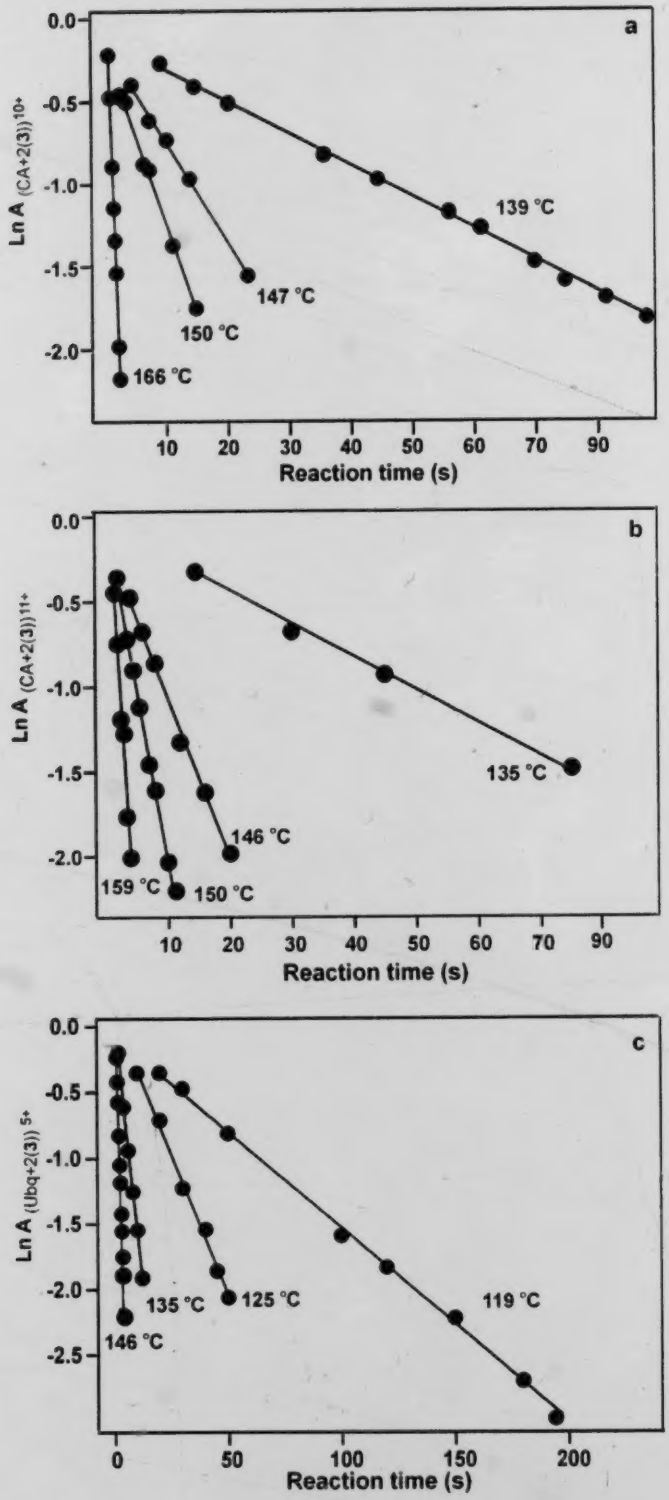

a
Ln A (CA+2(3))10+
0.0
-0.5
-1.0
-1.5
-2.0
139 °C
147 °C
150 °C
166 °C
10
20
30
40
50
60
70
90
Reaction time (s)
b
Ln A (CA+2(3))11+
0.0
-0.5
-1.0
-1.5
-2.0
135 °C
146 °C
159 °C
150 °C
10
20
30
40
50
60
70
90
Reaction time (s)
c
Ln A (Ubq+2(3)) 5+
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
135 °C
125 °C
119 °C
146 °C
0
50
100
150
200
Reaction time (s)

Figure 5.6 Plots of the natural logarithm of the normalized abundance of the complex ion: (a) $(CA + 2(\mathbf{3}))^{10+}$, (b) $(CA + 2(\mathbf{3}))^{11+}$ and (c) $(Ubq + 2(\mathbf{3}))^{5+}$, at the temperatures indicated. Kinetic data fit to first order kinetics.

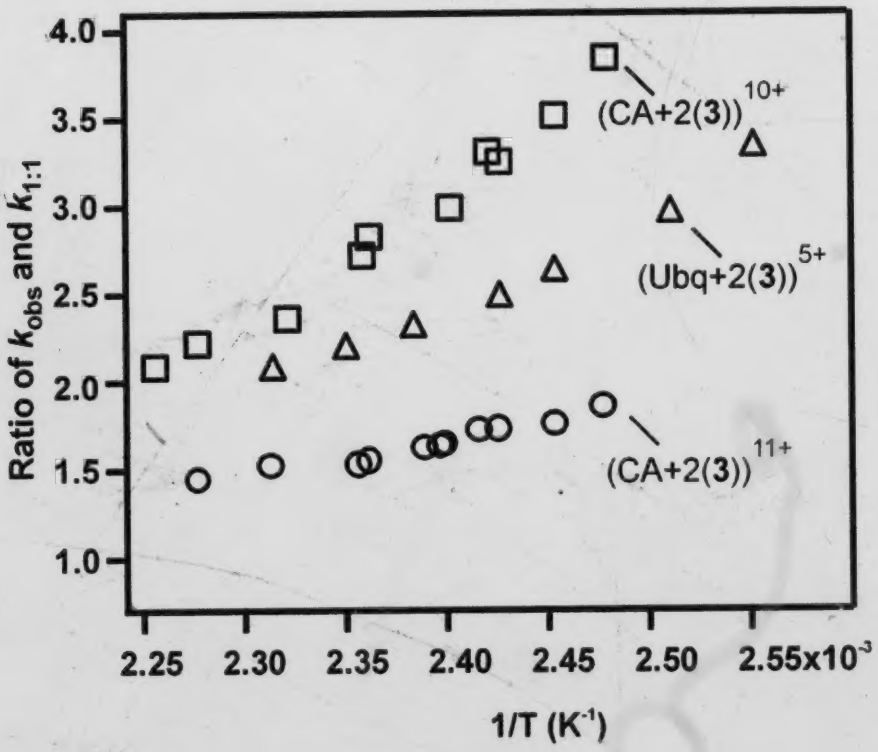


Figure 5.7 The ratio of the rate constant of the dissociation of 1:2 complex ion (k_{obs}) and the rate constant for the dissociation of corresponding 1:1 complex ion ($k_{1:1}$) at a given temperature (■, (CA + (2)**3**)$^{10+}$; ●, (CA + (2)**3**)$^{11+}$ and ▲, (Ubq + (2)**3**)$^{5+}$).

all based on reaction pathways described in Scheme 5.3, were considered. In *model I* the nonspecific binding sites are nonequivalent and independent, such that $k_1 = k_1' = k_{1:1}$ and $k_2 = k_2' = k_{obs} - k_{1:1}$. In the other two models, the binding sites are treated as equivalent but dependent; *model II*: $k_1 = k_2 = k_{obs}/2$ and $k_1' = k_2' = k_{1:1}$ and *model III*: $k_1 = k_2 = k_1' = k_2' = k_{obs}/2$. To determine whether any one of these models describes the dissociation kineticsmeasured experimentally, the breakdown curves simulated from these kinetic models were compared with the kinetic data measured for the 1:2 complexes, (CA +

$2(\mathbf{3}))^{10+}$, $(CA + 2(\mathbf{3}))^{11+}$ and $(Ubq + 2(\mathbf{3}))^{5+}$. Because the ratio $k_{obs}/k_{1:1}$ measured for the complexes $(CA + 2(\mathbf{3}))^{10+}$ and $(Ubq + 2(\mathbf{3}))^{5+}$ at high reaction temperatures and for $(CA + 2(\mathbf{3}))^{11+}$ at low reaction temperatures are close to 2, the values of $k_{obs}/2$, $k_{obs} - k_{1:1}$, and $k_{1:1}$ are similar making it difficult to distinguish between the theoretical curves derived from the different kinetic models. Therefore, the comparison of the theoretical and experimental breakdown curves were performed on the kinetic data obtained at a relatively low reaction temperature for the $(CA + 2(\mathbf{3}))^{10+}$ and $(Ubq + 2(\mathbf{3}))^{5+}$ ions, of 139 °C and of 119 °C, respectively, and relatively high temperature for $(CA + 2(\mathbf{3}))^{11+}$, 150 °C. Shown in Figure 5.8 are the theoretical and experimental breakdown curves for the three ions. It can be seen that the kinetic data measured for the $(CA + 2(\mathbf{3}))^{10+}$ and $(CA + 2(\mathbf{3}))^{11+}$ ions is best described by a model in which the two binding sites are equivalent but dependent. In the case of $(CA + 2(\mathbf{3}))^{10+}$, neither binding site resembles (kinetically) the interaction in the 1:1 complex, which means that the addition of the second carbohydrate changes irreversibly (at least on the time scale of the experiment) the protein-carbohydrate interactions, *i.e.* $k_1 = k_2 = k_1' = k_2' = k_{obs}/2$, *model III*. For $(CA + 2(\mathbf{3}))^{10+}$, after the loss of one carbohydrate, loss of the remaining carbohydrate resembles (kinetically) the interaction in the 1:1 complex, *i.e.* $k_1 = k_2 = k_{obs}/2$ and $k_1' = k_2' = k_{1:1}$, *model II*. In contrast to the results obtained for the $(CA + 2(\mathbf{3}))^{n+}$ ions, which indicate the presence of equivalent but dependent binding sites in the gas phase, the kinetic data measured for $(Ubq + 2(\mathbf{3}))^{5+}$ are best reproduced by *model I*, in which the binding sites are nonequivalent and independent, *i.e.* $k_1 = k_1' = k_{1:1}$ and $k_2 = k_2' = k_{obs} - k_{1:1}$. This intriguing result suggests that there exists a single, preferred interaction site in the gas phase. If one accepts that, at the time of their formation, the carbohydrate(s) in the 1:1

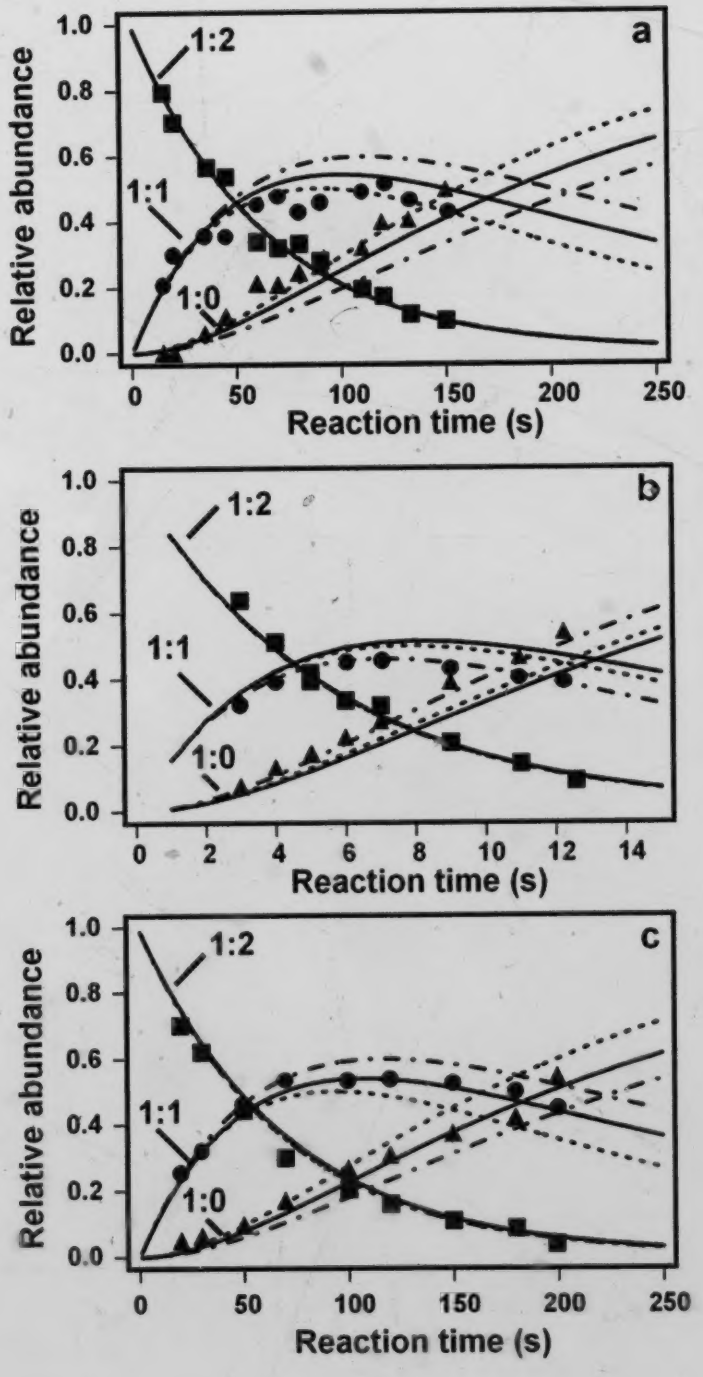

a
1:2
1:1
1:0
Relative abundance
Reaction time (s)
b
1:2
1:1
1:0
Relative abundance
Reaction time (s)
c
1:2
1:1
1:0
Relative abundance
Reaction time (s)

Figure 5.8 Comparisons of the experimental breakdown curves derived from the kinetic data for the complexes: (a) $(CA + 2(\mathbf{3}))^{10+}$ at 139 °C; (b) $(CA + 2(\mathbf{3}))^{11+}$ at 150 °C; (c) $(Ubq + 2(\mathbf{3}))^{5+}$ at 119 °C, and the corresponding theoretical breakdown curves calculated using the kinetic *model I* ($k_1 = k_1' = k_{1:1}$ and $k_2 = k_2' = k_{obs} - k_{1:1}$ (—)); kinetic *model II* ($k_1 = k_2 = 1/2k_{obs}$ and $k_1' = k_2' = k_{1:1}$ (.....); kinetic *model III* ($k_1 = k_2 = k_1' = k_2' = 1/2k_{obs}$ (— ·— ·)).

and 1:2 complexes samples a variety of sites on Ubq, then this result implies that the carbohydrates are able to rapidly migrate to the preferred binding sites. This result may have important implications for the retention of solution structure in gaseous protein-ligand complexes produced by ES (or nanoES) from specific interactions in solution.

5.3.4 Arrhenius Activation Parameters

Arrhenius plots were constructed from the temperature–dependent dissociation rate constants measured for the loss of neutral ligand from the 1:1 nonspecific complexes composed of CA, Ubq or BPTI and one of the carbohydrates, **3** or **5** – **7** (Figures 5.9-5.11). The Arrhenius activation energy, E_a, was determined from the slope of a linear least-squares fit of the Arrhenius plot and the preexponential (A) factor was determined from the y-intercept. The Arrhenius activation parameters and the corresponding entropy of activation ($\Delta S^{\ddagger}$) which were calculated at 415 K using eq 5.6, are listed in Table 5.1. Arrhenius plots could not be constructed for the complexes of CA and Ubq with **1** and **2**, because of the difficulty in extracting dissociation rate constants, and with **4** due to the limited range of reaction temperatures that could be investigated.

$$A = (ek_BT/h)\exp(\Delta S^{\ddagger}/R) \qquad (5.6)$$

Analysis of the Arrhenius parameters reveals that the stability of the nonspecific complexes is sensitive to a number of factors: structure of the carbohydrate and protein and the charge state of the complex. Each of these factors and their influence on the stability of the complexes are discussed below.

Table 5.1 Arrhenius parameters determined for the dissociation of gaseous, protonated and deprotonated protein-trisaccharide complexes: $(P + L)^{n+/z-} \rightarrow P^{n+/z-} + L$, where L = αTal[αAbe]αMan **(3)**, αGal[αAbe]αMan **(5)**, αGal[αAbe](4-doxyαMan) **(6)**, (6-deoxyαGal)[αAbe]αMan **(7)**.

P	L	n^+/z^-	E_a [a] (kcal/mol)	A [a] ($s^{\circ l}$)	ΔS° [b] (cal/mol·K)
CA	**3**	+8	51.3 ± 1.8	$10^{25.5 \pm 0.9}$	56
	3	+9	56.9 ± 2.1	$10^{28.2 \pm 0.9}$	68
	3	+10	61.2 ± 0.9	$10^{30.2 \pm 0.5}$	77
	3	+11	51.0 ± 0.6	$10^{25.4 \pm 0.3}$	55
	3	+12	49.5 ± 0.7	$10^{24.8 \pm 0.4}$	52
	5	+10	59.7 ± 1.1[c]	$10^{29.5 \pm 0.5c}$	74[c]
	5	+11	54.0 ± 0.5[c]	$10^{27.0 \pm 0.3c}$	62[c]
	6	+10	54.1 ± 0.8	$10^{27.1 \pm 0.4}$	63
	6	+11	52.4 ± 0.6	$10^{26.5 \pm 0.3}$	60
	7	+10	52.2 ± 0.9	$10^{26.1 \pm 0.5}$	58
	7	+11	52.2 ± 0.7	$10^{26.3 \pm 0.4}$	59
Ubq	**3**	+5	47.8 ± 1.6	$10^{24.4 \pm 0.9}$	51
	3	-5	48.0 ± 0.7	$10^{23.8 \pm 0.4}$	48
	5	+5	46.8 ± 1.2	$10^{24.2 \pm 0.6}$	50
BPTI	**5**	+5	46.0 ± 0.7	$10^{22.6 \pm 0.4}$	42

a. The reported errors are values of one standard deviation. b. Values calcula[illegible] 415 K from the corresponding *A*-factors. c. Arrhenius parameters taken from Chapter 3.

5.3.4.1 Influence of Carbohydrate Structure

Previous studies of the thermal dissociation of gaseous protein-carbohydrate complexes originating from specific interactions in solution have shown that the energetic and kinetic stability of the complexes are strongly influenced by intermolecular hydrogen bonds (H-bonds) between the protein and carbohydrate hydroxyl groups [10,

11]. The nonspecific complexes are expected to be similarly stabilized and their stability should reflect the number of available H-bond donor/acceptor groups, in particular thenumber of hydroxyl groups. As described previously [10, 11], the contribution of individual carbohydrate OH groups to the stability of protein-carbohydrate complexes can be evaluated from the change in dissociation E_a upon deoxygenation of the carbohydrate at specific sites. The contribution of the Man C-4 and Gal C-6 OH groups to the energetic stability of the (CA + **5**)$^{n+}$ ions, where n = 10 and 11, was assessed from the differences in the values of E_a determined for the complex (CA + **5**)$^{n+}$ and the corresponding complexes of the two monodeoxy congeners of **5**, (CA + **6**)$^{n+}$ and (CA + **7**)$^{n+}$. Deoxygenation at either site results in a decrease in E_a of 6 (**6**) and 8 kcal/mol (**7**) at +10 and 2 kcal/mol (**6** and **7**) at +11. The decrease in E_a strongly suggests that both hydroxyl groups interact with CA in the gas phase. Because of conformational constraints, the Gal C-6 and Man C-4 OH groups of **5** are not able to simultaneously solvate a single charge group, such as a protonated amino group or protonated imidazole group, indicating that one or both of the OH groups must be involved in neutral intermolecular H-bonds. While this result does not preclude the contribution of strong ionic H-bonds (*i.e.* charge solvation) to the stability of the (CA + **5**)$^{n+}$ ions, it does indicate that neutral interactions can play a significant role in stabilizing nonspecific protein-trisaccharide complexes. To investigate whether the configuration of individual carbohydrate OH groups (axial or equitoral) can influence the energetic stability of the complexes, the E_a values measured for (CA + **5**)$^{10/11+}$ and (CA + **3**)$^{10/11+}$ ions were compared. The structures of **5** and **3** differ only in the configuration of the OH group at C-2 of the nonreducing monosaccharide residue. At the +10 charge state, the E_a values

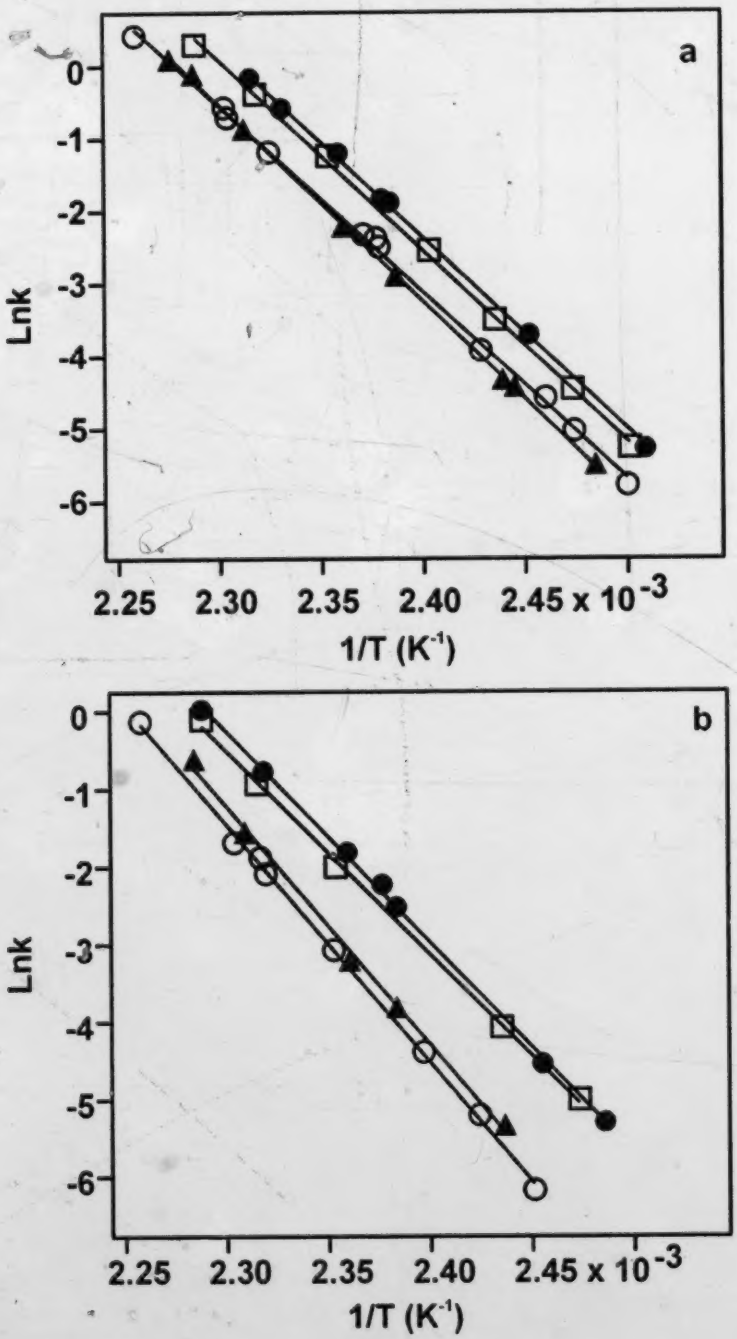


Figure 5.9 Arrhenius plots for the loss of L from the protonated complex ions, $(CA + L)^{n+}$ where n is (a) 10 and (b) 11, (○: L = **5**; ▲: L = **3**; ●: L = **6**; □: L = **7**).

are indistinguishable, within experimental uncertainty, suggesting that either the C-2 OH groups do not interact with the protein or that they engage in equivalent interactions. In contrast, at +11, there is a difference of 3 kcal/mol in the E_a values, a result that indicates a different energetic contribution from the C-2 OH group in these two complexes.

5.3.4.2 Influence of Protein Structure

The above analysis reveals that the energetic stability of the nonspecific (CA-trisaccharide)$^{n+}$ complexes is sensitive to the structure of the carbohydrate. There is also compelling evidence that the complexes are stabilized by multiple interactions, some of which are neutral H-bonds. Consequently, the structure of the protein and, in particular, the structure of the protein surface presented to the carbohydrate might also be expected to influence the energetic stability of the nonspecific complexes. However, in Chapter 3, it was shown that the E_a values for the dissociation of the protonated nonspecific 1:1 complexes of **5** with CA and with scFv, at the same charge state (+10 and +11), are indistinguishable, ~60 kcal/mol. This result suggests that protein structure (primary or higher order) does not significantly influence the energetic stability of the complex. This conclusion is further supported by similar dissociation E_a values (~47 kcal/mol) determined in the present work for the (BPTI + **5**)$^{5+}$ and (Ubq + **5**)$^{5+}$ ions, Table 5. An absence of a dependence of stability on protein structure for complexes at the same charge state could be interpreted to mean that the stabilizing intermolecular interactions are entirely ionic in nature. However, as described above, there is compelling evidence, in the case of the (CA + **5**)$^{10/11+}$ ions at least, that neutral interactions contribute to the

stability of the protein-trisaccharide complexes. Assuming that neutral intermolecular interactions are generally present in the nonspecific complexes, the present results imply that the bound carbohydrates are able to form energetically equivalent noncovalent interactions with the protein, independent of the nature of the functional groups on the surface of the protein. This very surprising result may point to a significant role played by the protein backbone in stabilizing the nonspecific complexes.

Although the protein structure does not influence the value of E_a for nonspecific complexes of the same charge state, it can influence kinetic stability. For example, the kinetic stability of the (BPTI + **5**)$^{5+}$ and (Ubq + **5**)$^{5+}$ ions are markedly different (at a given temperature), despite the similar E_a values. The difference in kinetic stability reflects a significant difference in $\Delta S^{\ddagger}$, with the more stable (BPTI + **5**)$^{5+}$ having a $\Delta S^{\ddagger}$ that is 8 cal/mol·K smaller than that of (Ubq + **5**)$^{5+}$. Differences in $\Delta S^{\ddagger}$ of similar magnitude have also been reported in Chapter 3 for the (CA + **5**)$^{10/11+}$ versus (scFv + **5**)$^{10/11+}$ ions. The origin of these differences in $\Delta S^{\ddagger}$ is not currently known and is the focus of ongoing research in our laboratory.

5.3.4.3 Influence of Charge State

To evaluate the influence of charge on the kinetic and energetic stability of the nonspecific complexes, the Arrhenius parameters determined for the (CA + **3**)$^{n+}$ ions at charge states ranging from +8 to +12 were compared. It can be seen from Figure 5.10 and Table 5.1 that the kinetic and energetic stability of the complex over the range of temperatures investigated is sensitive to its charge state; the trend in stability follows the

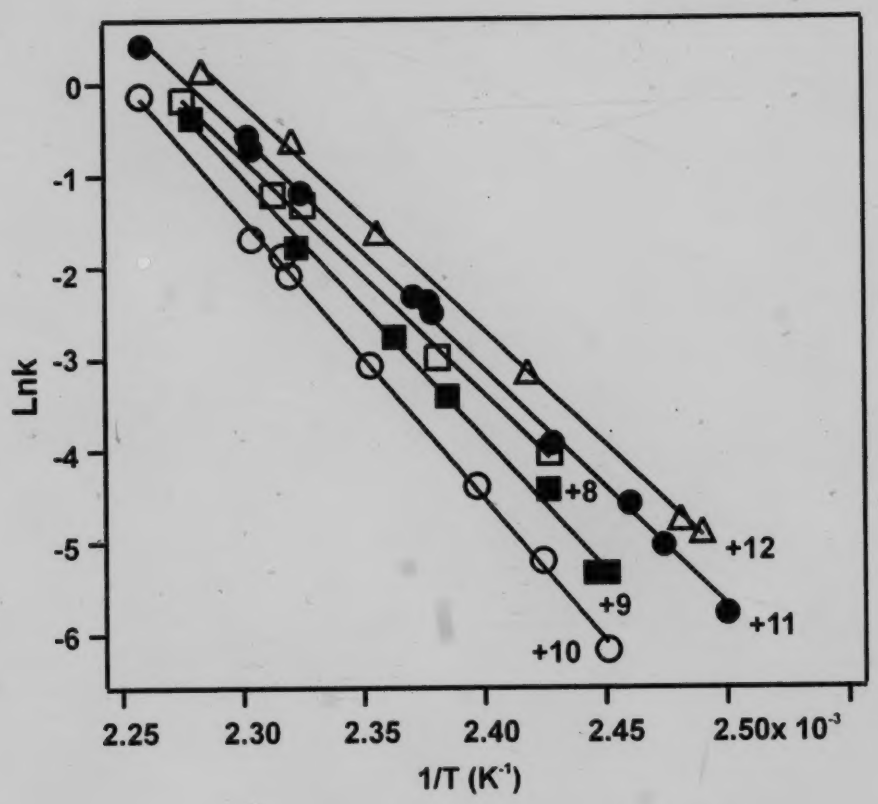


Figure 5.10 Arrhenius plots for the loss of **3** from the protonated complex ions, $(CA + 3)^{n+}$: □ (n = 8), ■ (9), ○ (10), ● (11), Δ (12).

order: +12 < +11 < +8 < +9 < +10. The significant decrease in E_a observed at charge states ≥ +10 is inconsistent with purely ionic intermolecular interactions. If binding was due solely to ionic interactions, E_a would be expected to increase with charge state due to the increasing acidity of protonated groups. The present results, therefore, likely reflect differences in the relative contribution of ionic and neutral interactions to the stability of the complex at the different charge states. Ion mobility measurements performed on a number of gaseous proteins produced by ES have shown that the protein ions adopt

compact structures at low charge states and more extended or unfolded structures at higher charge states [18]. It is reasonable, therefore, to assume that, of the charge states investigated, CA is in its most compact structure at +8 and that it adopts a more extended structure at higher charge states. Because proteins in more compact structures are able to more efficiently solvate charge [18e], ionic interactions between the CA and **3** are expected to be weakest at +8 and increase with increasing charge. In contrast, the contribution of neutral interactions between CA and **3** is expected to diminish as the protein unfolds at higher charge state. Because of the opposite trends for the strength of the ionic and neutral intermolecular interactions with charge state, the stability of the nonspecific complex is expected to initially increase with charge and then decrease at higher charge states, in agreement with the observed trend in the E_a values.

Additional support for the dominant influence of neutral intermolecular interactions at low charge states can be found in the similar E_a values (48 kcal/mol) measured for the $(Ubq + \mathbf{3})^{5+}$ and $(Ubq + \mathbf{3})^{5-}$ ions, Table 5.1. Because of differences in the nature and location of the charge groups in the protonated and deprotonated ions, the energetic stabilities of the $(Ubq + \mathbf{3})^{5+}$ and $(Ubq + \mathbf{3})^{5-}$ ions would be expected to differ if ionic interactions are present. According to thermochemical data reported for the sequential hydration of $n\text{-}C_3H_7NH_3^+$ and CH_3COO^- ions ($n\text{-}C_3H_7NH_3^+$: $\Delta H^\circ_{0,1}$ = 15.1 kcal/mol, $\Delta H^\circ_{1,2}$ = 11.6; CH_3COO^-: $\Delta H^\circ_{0,1}$ =17.1, $\Delta H^\circ_{1,2}$ = 12.8 [19]), which are the best available model systems for ionic H-bonding between carbohydrates and protonated and deprotonated amino acids, stronger intermolecular interactions are expected in the case of the deprotonated complex (assuming the same degree of intramolecular charge solvation in the two cases). The similarity in the E_a values determined for the $(Ubq + \mathbf{3})^{5+/5-}$ ions is,

therefore, inconsistent with the presence of ionic interactions. Instead, this result strongly suggest that the complexes are stabilized entirely by neutral intermolecular interactions. This conclusion is also consistent with the kinetic data obtained for the $(Ubq + 2(\mathbf{3}))^{5+}$ ion which indicate the presence of single carbohydrate binding site, in contrast to multiple equivalent sites for the $(CA + 2(\mathbf{5}))^{10/11+}$ ions wherein both ionic and neutral interactions are involved in stabilization.

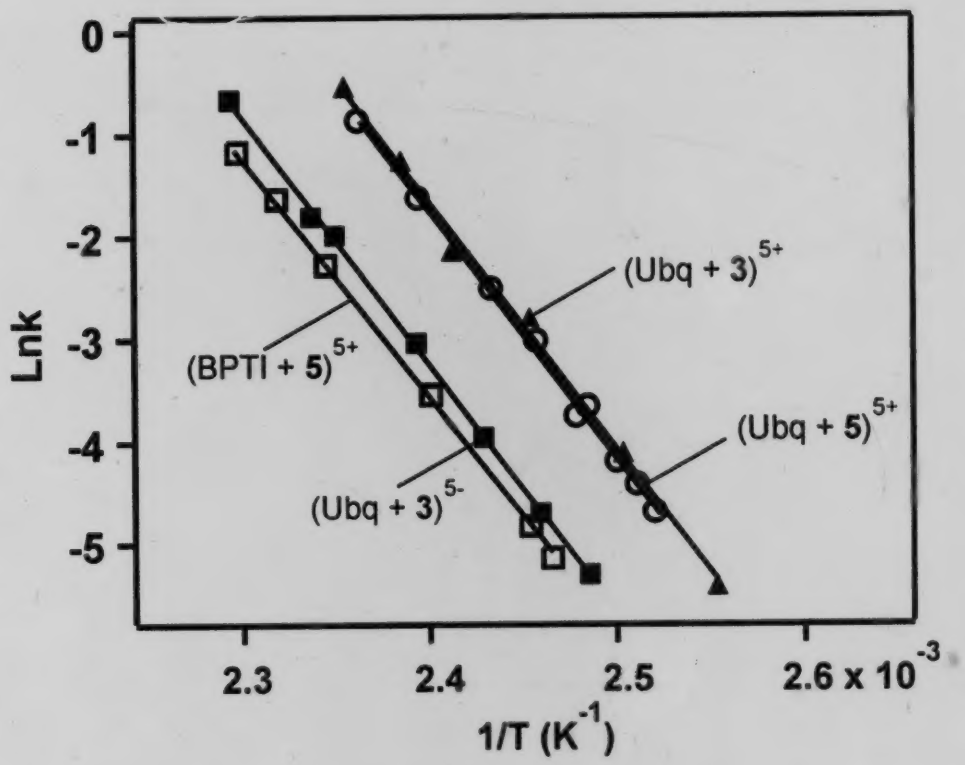


Figure 5.11 Arrhenius plots for the loss of L from the protonated and deprotonated complex ions (○, $(Ubq + \mathbf{5})^{5+}$; ▲, $(Ubq + \mathbf{3})^{5+}$; ■, $(Ubq + \mathbf{3})^{5-}$ and □, $(BPTI + \mathbf{5})^{5+}$).

5.4 Conclusions

Time-resolved thermal dissociation experiments, implemented with the BIRD technique and FT-ICR/MS, have been performed on a series of gaseous protonated and deprotonated protein-carbohydrate complexes produced by nonspecific interactions during the nanoES process. The results of this study have provided new insight into the nature of the intermolecular interactions in nonspecific protein-carbohydrate complexes in the gas phase. Over the range of temperatures investigated, the complexes were found to dissociate exclusively by the loss of the carbohydrate in its neutral form. Kinetic data measured for the dissociation of the 1:1 complexes containing a mono- or disaccharide suggest that multiple kinetically distinct complex structures are produced by nanoES. This result is consistent with the carbohydrates binding at multiple sites on the proteins in the gas phase. In contrast, kinetic data measured for the 1:1 complexes consisting of a tri- or tetrasaccharide suggest that the gaseous proteins have a single carbohydrate binding site or multiple equivalent binding sites. Analysis of the kinetic data for the dissociation of several 1:2 protein-carbohydrate complexes revealed that Ubq presents a single trisaccharide binding site in the gas phase, while CA offers multiple equivalent, but dependent, sites of interaction.

Arrhenius activation parameters were determined for the dissociation of a series of 1:1 complexes of CA with structurally-related trisaccharides. Compelling evidence that neutral intermolecular hydrogen bonds contribute to the stability of the complexes, at least at certain charge states, was obtained. Surprisingly, the energetic stability of the complexes was found to be insensitive to the structure of the protein, suggesting that the carbohydrates are able to form energetically equivalent interactions with the various

functional groups presented by the protein. Although the protein structure does not influence the magnitude of the dissociation E_a it can significantly influence the magnitude of $\Delta S^{\ddagger}$. However, the origin of this entropy difference is unclear. The energetic stability of a given protein-trisaccharide complex was found to be sensitive to its charge state, although no simple relationship between E_a and charge state is evident. It is proposed that both ionic and neutral hydrogen bonds are responsible for stabilizing nonspecific protein-carbohydrate complexes in the gas phase and that the relative contribution of neutral and ionic interactions is strongly influenced by charge state, with neutral interactions dominating at low charge states and ionic interactions dominating at high charge states.

5.5 Literature Cited

(1) Smith D. L.; Zhang Z. *Mass Spectrom. Rev.* **1994**, *13*, 411-429.

(2) Smith R. D.; Bruce J. E.; Wu Q.; Lei Q. P. *Chem. Soc. Rev.* **1997**, *26*, 191-202.

(3) Loo J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.

(4) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(5) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1-27.

(6) Heck, A. J. R.; van den Heuvel, R. H. H. *Mass Spectrom. Rev.* **2004**, *23*, 368-389.

(7) Li, Y. T.; Hsieh, Y. L.; Henion, J. D.; Senko, M. W.; McLafferty, F. W.; Ganem, B. *J. Am. Chem. Soc.* **1993**, *115*, 8409-8413.

(8) Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. *J. Am. Chem. Soc.* **1996**, *118*, 8646-8653.

(9) Wu, Q.; Gao, J.; Joseph-McCarthy, D.; Sigal, G. B.; Bruce, J. E.; Whitesides, G. M.; Smith, R. D. *J. Am. Chem. Soc.* **1997,** *119*, 1157-1158.

(10) Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002,** *124*, 5902-5913.

(11) Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002,** *124*, 9340-9341.

(12) Felitsyn, N; Peschke, M.; Kebarle, P. *Int. J. Mass Spectrom.* **2002**, *219*, 39-62.

(13) (a) Carpenter, J. F.; Crowe, J. H. . *Biochemistry* **1989**, *28*, 3916-3922; (b) Murray, B. S.; Liang, H.-J. *Langmuir* **2000**, *16*, 6061-6063.

(14) Felitsyn, N; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2001**, *73*, 4647-4661.

(15) Hunter, E. P. L.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, 27, 413-656.

(16) (a) Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194-197; (b) Price, W. D.; Williams, E. R. *J. Phys. Chem. A* **1997**, *101*, 8844-8852.

(17) Price, W. D.; Williams, E. R. *J. Phys. Chem. B* **1997**, *101*, 664-673.

(18) (a) Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1995**, *117*, 10141-10142; (b) Shelimov, K. B.; Jarrold, M. F. *J. Am. Chem. Soc.* **1996**, *118*, 10313-10314; (c) Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 2240-2248; (d) Fye, J. L.; Woenckhaus, J.; Jarrold, M. F. *J. Am. Chem. Soc.* **1998**, *120*, 1327-1328; (e)Li, J.; Taraszka, J. A.; Counterman, A. E.; Clemmer, D. E. *Int. J. Mass Spectrom.* **1999**, *185/186/187*, 37-47; (f) Mao, Y.; Ratner, M. A.; Jarrold, M. F. *J. Phys. Chem. B.* **1999**, *103*, 10017-10021.

(19) (a) Meot-Ner, M. *J. Am. Chem. Soc.* **1984**, *106*, 1265-1272; (b) Meot-Ner, M. *J. Am. Chem. Soc.* **1988**, *110*, 3854-3858.

Chapter 6

Conclusions and Future Work

The objectives of this thesis work were to develop MS–based approaches for studying the affinity and specificity of protein-carbohydrate interactions. Research work described in this thesis can be categorized into two main aspects:

(1) Determination of protein-carbohydrate binding affinities by nanoES-MS. Particular effort was devoted to investigating the influence of the solution and gas phase processes on binding affinity measurements. (Chapter 2).

(2) Investigation of the formation of nonspecific protein-carbohydrate complexes by nanoES and their structures and stabilities in the gas phase. This study demonstrated the intrinsic interactions of gaseous complexes (Chapters 3, 4, and 5).

In Chapter 2, an extensively studied complex, single-chain variable domain fragment (scFv) of a carbohydrate–binding antibody and its native trisaccharide served as a model system to investigate the effects of spray duration, analyte concentration and ion source conditions on the measurement of the binding constant by using nanoES-FT-ICR/MS. The experimental conditions that preserve the original solution composition throughout the formation of gaseous ions have been identified. Under these optimized experimental conditions, binding constants of a series of protein-carbohydrate complexes were accurately measured. Beyond the determination of binding affinities of complexes, factors that significantly influence the relative intensities of the bound and free protein observed in mass spectra and, hence, the binding constant measurements, were investigated. These factors include nonspecific association between protein and

carbohydrate molecules during the nanoES, solution pH that changes with spray duration and in-source dissociation. Studies on these factors are crucial in helping us understand how the solution and gas phase processes influence the binding affinity measurements using ES-MS method. This understanding is essential to developing MS-based methods for reliable and accurate quantitative determination of binding affinities, particularly for weakly bound complexes. A possible extension of this study is to apply the developed nanoES-MS approach to determine the binding constants for other classes of noncovalent complexes. For instance, it has been shown that the binding affinities measured by ES-MS for the protein-acyl CoA complex [1] and ribonuclease A-cytidine 2'-monophosphate complex [2] did not correlate with their solution affinities. Therefore, future work will focus on expanding and developing nanoES-MS approaches for the binding constant measurements on such complexes.

The tendency to form nonspecific protein-carbohydrate complexes by nanoES was found to be pronounced at high carbohydrate concentrations, which led to artificially high binding affinities measured by the direct nanoES/MS technique. It is therefore desirable to eliminate nonspecific complexes prior to MS detection. An approach of selectively dissociating nonspecific complexes in the gas phase has been proposed. [3] To examine the feasibility of this approach, a comparative study of the kinetic and energetic stabilities of gaseous specific and nonspecific protein-carbohydrate complexes was carried out in Chapter 3. The results from thermal dissociation experiments using BIRD technique revealed that the protonated nonspecific protein-carbohydrate complex is kinetically more stable than the corresponding specific complex at +10 and +11 charge states. More importantly, for the +10 ions, the nonspecific complex is energetically more

favorable. This study clearly indicated that gas phase dissociation can not lead to the selective removal of nonspecific protein-carbohydrate complexes from specific complexes. In order to establish a strategy to minimize the inference of nonspecific complexes, a better understanding of the mechanism of formation and the nature of the interactions stabilizing the nonspecific complexes in the gas phase will be vital. Therefore, nonspecific protein-carbohydrate binding was investigated in more detail in Chapters 4 and 5. In Chapter 4, by performing the titration experiments on a series of protein-carbohydrate nonspecific complexes, we concluded that nonspecific binding has dependence on the size and hydrophilicity of carbohydrates, but not on the size of proteins. It was also found that the efficiency of formation of nonspecific complexes is greater for small or hydrophilic carbohydrates than for larger or more hydrophobic complexes. In addition, thermal dissociation experiments performed on the gaseous nonspecific complexes shown in Chapter 4 reveal that their kinetic stability increases with the size of carbohydrates and proteins. Based on the trend in gas phase stability of the complexes, the number of carbohydrate ligands bound to protein is expected to increase with the ligand size. However, experimental results are opposite to this prediction. It suggests that gas phase dissociation of complexes prior to detection is not responsible for the smaller number of carbohydrate ligands bound to protein observed for larger carbohydrate ligands. We propose that the efficiency of the nonspecific binding process and the distribution of bound carbohydrates is governed, predominantly, by the size of the offspring droplets that ultimately produce the gaseous ions. Moreover, the number of carbohydrate ligands bound to the protein is influenced by the loss of carbohydrate molecules from the offspring droplets through ion evaporation and the self-

association of carbohydrates into clusters within the offspring droplets. Based on the competitive nature of gaseous carbohydrate ion formation and the formation of the nonspecific protein-carbohydrate complexes by nanoES, a new strategy, which employs the use of metal salts to minimize the formation of nonspecific protein-carbohydrate complexes during the nanoES process, is proposed. The addition of Na^+, Ca^{2+} and Mg^{2+} to the nanoES solution was shown to enhance the formation of the gaseous carbohydrate ions, either as monomers or clusters, and reduce the degree of nonspecific protein-carbohydrate binding. The divalent Ca^{2+} and Mg^{2+} ions were found to be more effective than the monovalent Na^+, although the effectiveness of all of the salts increased with concentration.

Further work will involve the application of this strategy to a specific protein-carbohydrate complex system. Our laboratory has applied nanoES-FT-ICR/MS to investigate the binding of the P^k trisaccharide to the multivalent protein complex SLT-1 (B_5) [4]. Since the binding constant of this complex is low, at 10^3 M^{-1}, the use of high carbohydrate concentration is generally required in titration experiments. Since the formation of nonspecific complexes leads to a higher binding constant measured in MS, the strategy of using metal salt to minimize the nonspecific complexes will be applied to this system to examine the effectiveness of the developed strategy. An alternative approach is to use a protein that cannot specifically interact with P^k in solution as an internal reference for the determination of the fraction of B_5 engaged in nonspecific interactions with P^k. Since the degree of nonspecific binding was found to be insensitive to the structure of protein, as shown in Chapter 4, the amount of nonspecific complexes formed between B_5 and its specific ligand P^k is expected to be equal to the amount of

nonspecific complexes formed between P^k and the internal reference protein. The amount of nonspecific complexes of P^k with the internal reference protein can be easily obtained based on the abundance ions observed in mass spectra. Thus, the amount of specific complexes of B_5 and P^k can be deduced from the total amount of complexes observed in mass spectra. Using this correction, the binding constant of this complex is expected to be measured more accurately by MS.

Also in Chapter 4, we noticed that the number of bound carbohydrate ligands exceeds the charge states of the complex, which indicates that the neutral interactions play a role in the formation of nonspecific complexes. A detailed study of the nature of nonspecific binding was further carried out in Chapter 5. Gas phase thermal dissociation experiments using BIRD technique was performed on a series of 1:1 and 1:2 nonspecific protein-carbohydrate complexes. The results from this study revealed that the 1:1 complex containing mono- or disaccharide has multiple kinetically distinct structures, while the complex containing tri- or tetrasaccharide has a single favorable binding structure or multiple equivalent binding structures. Analysis of the Arrhenius activation parameters (E_a and A) determined for the dissociation of 1:1 protein-trisaccharide complexes demonstrated that the kinetic and energetic stability of the complexes has dependence on the structure of carbohydrate and the charge state of the complex. The energetic stability of nonspecific complexes was found to be sensitive to the size of protein compared to the protein structure (primary or higher order). We propose that both ionic and neutral hydrogen bonds are responsible for stabilizing nonspecific protein-carbohydrate complexes in the gas phase and that the relative contribution of the neutral and ionic interactions is strongly influenced by charge state.

Literature Cited

(1) Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. *J. Am. Chem. Soc.* **1996**, *118*, 8646-8653.

(2) Benkestock, K.; Sundqvist, G.; Edlund, P.-O.; Roeraade, J. *J. Mass Spectrom.* **2004**, *39*, 1059-1067.

(3) Smith R. D.; Bruce J. E.; Wu Q.; Lei Q. P. *Chem. Soc. Rev.* **1997**, *26*, 191-202.

(4) (a) Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. Glycobiology 2001, 11, 605-611; (b) Daneshfar, R; Kitova, E. N.; Klassen, J. S. *J. Am. Chem. Soc.* **2004**, *126*, 4786-4787.